

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 1, 2003, 07:49:49 ; Search time 39.3958 Seconds
(without alignments)
412.647 Million cell updates/sec

Title: US-09-987-485-1

Perfect score: 615

Sequence: 1 MKLKVTNVTAYDVIDVDK.....KVLKRDVAVGGGGLIKIG 2722

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : A_Geneseq.101002.*
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	615	100.0	126	21	AA1984002
2	615	100.0	194	21	AA178908
3	615	100.0	253	18	AA111886
4	615	100.0	256	21	AA194001
5	615	100.0	266	21	AA194003
6	604.5	98.3	123	13	AA128177
7	436	70.9	125	22	AA151122
8	231.5	37.6	148	22	AA196610
9	193.5	31.5	100	14	AA194210
10	174	28.3	591	22	AA183169

11	173.5	28.2	134	22	AAU03126	Streptococcus pyog
12	170.5	27.7	591	22	AA193186	C glutamicum prote
13	170.5	27.7	591	22	AA193186	Corynebacterium gl
14	166	27.0	600	21	AA198784	M. tuberculosis an
15	162.5	26.4	1157	23	AA198050	Corynebacterium fe
16	162	26.3	1139	22	AA1983180	Corynebacterium th
17	159.5	25.9	272	22	AA1979298	Corynebacterium gl
18	159.5	25.9	272	22	AA1979298	Corynebacterium gl
19	159.5	25.9	1140	20	AA193971	C. glutamicum pyru
20	159.5	25.9	1140	22	AA193971	C. glutamicum prote
21	159.5	25.9	1140	22	AA193971	C. glutamicum prote
22	159.5	25.9	1140	22	AA193971	C. glutamicum prote
23	159.5	25.9	1140	23	AA198053	Corynebacterium gl
24	159.5	25.9	1141	21	AA198053	Corynebacterium wi
25	159	25.9	161	22	AA198053	Pyruvate carboxyla
26	159	25.9	161	22	AA198053	CFE 89 protein seq
27	157	25.5	163	19	AA198571	S. pneumoniae deri
28	157	25.5	161	22	AA198571	Streptococcus pneu
29	157	25.5	161	22	AA198571	Streptococcus pneu
30	156	25.4	1073	22	AA198571	Listeria monocytog
31	156	25.4	1073	22	AA198571	Staphylococcus aur
32	155.5	25.3	1140	23	AA198052	Staphylococcus aur
33	153	24.9	294	22	AA198052	Corynebacterium mu
34	153	24.9	377	22	AA198052	Novel human enzyme
35	153	24.9	725	22	AA198052	CoA carboxylase 41
36	153	24.9	725	22	AA198052	Human protein sequ
37	153	24.9	725	23	AA198052	DNA encoding human
38	151	24.6	166	23	AA198052	Human biotin enzym
39	150	24.4	108	20	AA198052	Streptococcus poly
40	150	24.4	108	20	AA198052	M. vaccae antigen
41	150	24.4	108	23	AA198052	M. vaccae GV-29 par
42	150	24.4	155	23	AA198052	Lactococcus lactis
43	149	24.2	1151	22	AA198052	Amino acid sequenc
44	149	24.2	1151	22	AA198052	S. epidermidis ope
45	148.5	24.1	582	22	AA198052	Staphylococcus epi
46	147	23.9	117	20	AA198052	Micromonospora eve
47	144	23.4	1142	22	AA198052	Human endometrium
48	143.5	23.3	166	23	AA198052	Enterococcus faeca
49	143.5	23.3	184	23	AA198052	Streptococcus poly
50	143.5	23.3	1136	23	AA198052	Streptococcus poly
51	139.5	22.7	160	22	AA198052	Lactococcus lactis
52	139.5	22.7	160	22	AA198052	Lactococcus lactis
53	136.5	22.2	158	17	AA198052	Enterococcus faeca
54	136.5	22.2	158	17	AA198052	Enterococcus faeca
55	134.5	21.9	1181	22	AA198052	Synechococcus blot
56	134.5	21.9	1181	22	AA198052	Synechococcus blot
57	134.5	21.9	1181	22	AA198052	Drosophila melanog
58	134.5	21.9	1181	22	AA198052	Drosophila melanog
59	133	21.6	1148	22	AA198052	Drosophila melanog
60	131	21.3	164	20	AA198052	Bacillus subtilis
61	130	21.1	630	22	AA198052	Protein involved i
62	130	21.1	630	22	AA198052	E. coli cellular p
63	129.5	21.1	182	17	AA198052	Novel mar regulate
64	129.5	21.1	182	19	AA198052	Anabaena biotin ca
65	127	20.7	590	23	AA198052	Anabaena biotin ca
66	127	20.7	590	23	AA198052	S. coelicolor Acca
67	123	20.0	74	22	AA198052	S. coelicolor Acca
68	121.5	19.8	101	23	AA198052	Novel human secret
69	121.5	19.8	154	22	AA198052	Human biotin-depen
70	120	19.5	154	22	AA198052	Staphylococcus aur
71	119.5	19.4	156	23	AA198052	Staphylococcus aur
72	119	19.3	619	22	AA198052	Staphylococcus epi
73	117.5	19.1	155	22	AA198052	Propionibacterium
74	117	19.0	156	16	AA198052	S. epidermidis ope
75	117	19.0	156	16	AA198052	S. epidermidis ope
76	117	19.0	1830	14	AA198052	Pseudomonas aerugi
77	115	18.7	71	19	AA198052	Urea amide lyase.
78	115	18.7	71	19	AA198052	Mycobacterium tube
79	115	18.7	111	23	AA198052	Amino acid sequenc
80	114.5	18.6	675	22	AA198052	Human carboxylase-
81	113.5	18.5	194	22	AA198052	C glutamicum prote
82	113.5	18.5	194	22	AA198052	Corynebacterium gl
83	112	18.2	155	22	AA198052	Corynebacterium gl
						Haemophilus influ

84 110 17.9 76 22 AAG75011 Human colon cancer
85 110 17.9 76 23 ABP41119 Human ovarian anti
86 110 17.9 720 22 ABG11227 Novel human diagno
87 110 17.9 728 22 AAM39515 Human polypeptide
88 110 17.9 734 22 AAM41301 Human polypeptide
89 109 17.7 22 16 AAR65021 Biotinylation pept
90 109 17.7 22 16 AAE09858 Propionibacterium
91 108.5 17.6 155 23 ABB48312 Listeria monocytog
92 108.5 17.6 167 20 AAY34780 Chlamydia pneumoni
93 108 17.6 146 22 AAG81854 S. epidermidis ope
94 108 17.6 146 23 ABP38749 Staphylococcus epi
95 105.5 17.2 613 22 AAU62343 Propionibacterium
96 104 16.9 362 21 AAG17822 Arabidopsis thalia
97 104 16.9 364 21 AAG46637 Arabidopsis thalia
98 104 16.9 461 21 AAG17821 Arabidopsis thalia
99 104 16.9 462 21 AAG17820 Arabidopsis thalia
100 104 16.9 463 21 AAG46636 Arabidopsis thalia

ALIGNMENTS

RESULT 1
AAY84002
ID AAY84002 standard; Protein; 126 AA.

AC AAY84002;
DT 03-JUL-2000 (first entry)
DE Amino acid sequence of a T. pallidum protein.
KW PinPoint vector; fusion protein antigen; membrane antigen; syphilis.

OS Treponema pallidum.

PN **EF98931-22**

PD 15-MAR-2000.

PF 12-AUG-1999; 99EP-0115877.

PR 04-SEP-1998; 98US-0148920.

PA (BECT) BECTON DICKINSON & CO.

PI **Treponema pallidum, Deutsche Str.**

DR WPI: 2000-226057/20.

DR N-PSDB; AAZ99247.

PT Immunoassay for anti-Treponema pallidum antibodies, used for diagnosis

PT of syphilis using a fusion protein of membrane antigen with peptide

PT sequence that can be biotinylated in vivo

PS Disclosure; Page 12-13; 16pp; English.

XX The present sequence represents a Treponema pallidum protein, which is

XX used to produce a fusion protein antigen for use in the method of

XX the invention. The specification describes a method for detecting

XX antibodies against Treponema pallidum. The antibodies are detected

XX in a sample by reaction with a fusion protein antigen, present in

XX the mixture in limiting concentration. The fusion protein antigen

XX comprises a Treponema pallidum membrane antigen. The method is used

XX for diagnosis of syphilis.

XX Sequence 126 AA;

SQ Query Match 100.0%; Score 615; DB 21; Length 126;

Best Local Similarity 100.0%; Pred. No. 2.4e-52;

Matches 122; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKLKVTNCTAVDVEDVDKSHENPMGTILFGGGTGAPAPAGGAGKAGEIPAPL 60

Db 1 MKLKVTNCTAVDVEDVDKSHENPMGTILFGGGTGAPAPAGGAGKAGEIPAPL 60
QY 61 AGTVSKILVKEGDTKAGQTVLVLEAMKMETEINAPTDGKVKVLYKRDVAVGGGGLIK 120
Db 61 AGTVSKILVKEGDTKAGQTVLVLEAMKMETEINAPTDGKVKVLYKRDVAVGGGGLIK 120
QY 121 IG 122
Db 121 IG 122

RESULT 2

AAY78908

ID AAY78908 standard; Protein; 194 AA.

AC AAY78908;

XX 19-MAY-2000 (first entry)

DE Outer surface protein C (OspC) DraI fragment amino acid sequence.

KW Outer surface protein C; OspC; immunological epitope; Lyme disease;

KW vaccine; prevention; Borrelia infection; diagnose.

OS Borrelia burgdorferi.

PN WO200006745-A1.

PD 10-FEB-2000.

PF 30-JUL-1999; 99WO-US17270.

PR 31-JUL-1998; 98US-0094955.

PA (GUND-) GUNDERSEN LUTHERAN MEDICAL FOUND INC.

PI Callister SN, Lovrich SD, Schell RF, Jobe DA;

DR WPI: 2000-195305/17.

DR N-PSDB; AAZ92216.

PT New immunogenic polypeptides useful as a vaccine against Lyme disease

PT and for treating and detecting borrelia infection in mammals consists

PT an epitope of Borrelia burgdorferi OspC fragment

PS Claim 3; Fig 4; 51pp; English.

XX This sequence represents the Borrelia burgdorferi outer surface protein C

XX (OspC) DraI fragment amino acid sequence. The polypeptide contains an

XX immunological epitope used in the invention. Large amounts of OspC are

XX rapidly synthesised by B. burgdorferi shortly after attachment of

XX infected ticks to mammalian hosts. The OspC protein sequence is used to

XX diagnose B. borrelia infection in mammals. The OspC nucleotide sequence

XX is used to prevent (via vaccination), treat or detect Borrelia

XX (especially B. burgdorferi) infections, i.e. Lyme disease, in mammals

XX including humans. The OspC nucleotide sequence provides a superior

XX diagnostic antigen that detects early Lyme disease infection, predicts

XX successful eradication or the organism from the host, and discriminates

XX between individuals with Lyme disease and individuals who have been

XX vaccinated with an OspA Lyme disease vaccine. Detection of anti-OspC

XX borrelial antibodies advantageously gives an early diagnosis which

XX anti-OspA and anti-OspB borrelial antibodies cannot do. Unlike

XX vaccination with OspA, vaccination with OspC results in clearance of

XX spirochetes and resolution of symptoms even if administered after

XX infection with B. burgdorferi.

SQ Query Match 100.0%; Score 615; DB 21; Length 194;

Best Local Similarity 100.0%; Pred. No. 4.1e-52;

Matches 122; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKLKVTNCTAVDVEDVDKSHENPMGTILFGGGTGAPAPAGGAGKAGEIPAPL 60

QY 1 MKLKTAVNGTAVDVVDVVDKSHENPMGTILFEGGTGAPAPAPAGAGAGKAGEGEIPAPL 60
 DB 1 MKLKTAVNGTAVDVVDVVDKSHENPMGTILFEGGTGAPAPAPAGAGAGKAGEGEIPAPL 60
 QY 61 AGTAVSKILVKEGDDTVKAGQTVLVLEAMKMETEINAPTDGKVEKVLVKERDAVGGGGLIK 120
 DB 61 AGTAVSKILVKEGDDTVKAGQTVLVLEAMKMETEINAPTDGKVEKVLVKERDAVGGGGLIK 120
 QY 121 IG 122
 DB 121 IG 122

RESULT 3
 AAM11886
 ID AAM11886 standard; protein: 253 AA.

AC AAM11886;
 DT 21-APR-1997 (first entry)
 XX Hantavirus nuclear protein sequence.
 DE Hantavirus nuclear protein sequence.
 XX Hantavirus nuclear protein: antigen; monoclonal antibody;
 KM diagnosis; specificity.
 OS Hantavirus.
 XX JF08325291-A.

XX 10-DEC-1996.
 PD 30-MAY-1995; 95JP-0132460.
 PF 30-MAY-1995; 95JP-0132460.
 PR 30-MAY-1995; 95JP-0132460.
 XX (AAT-) A & T KR.
 PA (ARIK/) ARIKAWA J.
 PA (HASH/) HASHIMOTO N.
 XX WPI: 1997-083468/08.

XX Hantavirus antigen protein and monoclonal antibody - used in the
 PT diagnosis and treatment of hantavirus infection
 PS Example 1; Page 17; 18pp; Japanese.

XX The sequences given in AAM11872-86 represent Hantavirus nuclear
 CC proteins. The N-terminal regions of these proteins may be used as
 CC antigens to raise anti-hantavirus monoclonal antibodies. These
 CC antibodies are useful in diagnosis of hantavirus infection due to
 CC their specificity to Hantavirus.

SO Sequence 253 AA;

Query Match 100.0%; Score 615; DB 18; Length 253;
 Best Local Similarity 100.0%; Pred. No. 5,7e-52;
 Matches 122; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKLKTAVNGTAVDVVDVVDKSHENPMGTILFEGGTGAPAPAPAGAGAGKAGEGEIPAPL 60
 DB 1 MKLKTAVNGTAVDVVDVVDKSHENPMGTILFEGGTGAPAPAPAGAGAGKAGEGEIPAPL 60
 QY 61 AGTAVSKILVKEGDDTVKAGQTVLVLEAMKMETEINAPTDGKVEKVLVKERDAVGGGGLIK 120
 DB 61 AGTAVSKILVKEGDDTVKAGQTVLVLEAMKMETEINAPTDGKVEKVLVKERDAVGGGGLIK 120
 QY 121 IG 122
 DB 121 IG 122

RESULT 4

AA84001
 ID AA84001 standard; Protein: 256 AA.
 AC AA84001;
 DT 03-JUL-2000 (first entry)
 XX Amino acid sequence of a T. pallidum fusion protein antigen.

XX PinPoint vector; fusion protein antigen; membrane antigen; syphilis.
 OS Synthetic.
 OS Treponema pallidum.

Ep985931-A2
 PD 15-MAR-2000.

XX 12-AUG-1999; 99EP-0115877.
 PF 04-SEP-1998; 98US-0148920.
 PR (BECT) BECTON DICKINSON & CO.
 PA Mullerix MC, Deutsch J;
 PI WPI: 2000-226057/20.
 DR N-PSDB; AA299246.

XX Immunoassay for anti-Treponema pallidum antibodies, used for diagnosis
 PT of syphilis using a fusion protein of membrane antigen with peptide
 PT sequence that can be biotinylated in vivo
 PS Claim 9; Page 10-11; 16pp; English.

XX The present sequence represents a Treponema pallidum fusion protein
 CC antigen. The protein is used in the method of the invention. The
 CC specification describes a method for detecting antibodies against
 CC Treponema pallidum. The antibodies are detected in a sample by
 CC reaction with a fusion protein antigen, present in the mixture in
 CC limiting concentration. The fusion protein antigen comprises a
 CC Treponema pallidum membrane antigen. The method is used for diagnosis
 CC of syphilis.

SO Sequence 256 AA;

Query Match 100.0%; Score 615; DB 21; Length 256;
 Best Local Similarity 100.0%; Pred. No. 5.8e-52;
 Matches 122; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKLKTAVNGTAVDVVDVVDKSHENPMGTILFEGGTGAPAPAPAGAGAGKAGEGEIPAPL 60
 DB 1 MKLKTAVNGTAVDVVDVVDKSHENPMGTILFEGGTGAPAPAPAGAGAGKAGEGEIPAPL 60
 QY 61 AGTAVSKILVKEGDDTVKAGQTVLVLEAMKMETEINAPTDGKVEKVLVKERDAVGGGGLIK 120
 DB 61 AGTAVSKILVKEGDDTVKAGQTVLVLEAMKMETEINAPTDGKVEKVLVKERDAVGGGGLIK 120
 QY 121 IG 122
 DB 121 IG 122

RESULT 5
 AA84003
 ID AA84003 standard; Protein: 266 AA.

XX AA84003;
 DT 03-JUL-2000 (first entry)

XX Amino acid sequence of a T. pallidum membrane protein antigen.

KW Pinpoint vector: fusion protein antigen; membrane antigen; syphilis.
 OS Treponema pallidum.
 XX EP985931-A2.
 PN 15-MAR-2000.
 XX 12-AUG-1999; 99EP-0115877.
 XX 04-SEP-1998; 98US-0148920.
 XX (BECT) BECTON DICKINSON & CO.
 XX Mullerix MC, Deusch J;
 XX WPI: 2000-226057/20.
 XX N-PSDB; AA299248.
 DR
 XX
 XX Immunassay for anti-Treponema pallidum antibodies, used for diagnosis
 PT of syphilis using a fusion protein of membrane antigen with peptide
 PT sequence that can be biotinylated in vivo
 PS
 XX Claim 9; Page 14-15; 16pp: English.

CC The present sequence represents a Treponema pallidum membrane protein
 CC antigen, which is used to produce a fusion protein antigen for use
 CC in the method of the invention. The specification describes a method
 CC for detecting antibodies against Treponema pallidum. The antibodies
 CC are detected in a sample by reaction with a fusion protein antigen,
 CC present in the mixture in limiting concentration. The fusion protein
 CC antigen comprises a Treponema pallidum membrane antigen. The method is
 CC used for diagnosis of syphilis.

SO Sequence 266 AA;

Query Match 100.0%; Score 615; DB 21; Length 266;

Best Local Similarity 100.0%; Pred. No. 6e-52; Mismatches 0; Indels 0; Gaps 0;

Matches 122; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKLKVYNGTAYDVVDVDSHENPMGTILFGGTCGAPAPAAAGAGAGAGEEIPAPL 60
 DB 1 MKLKVYNGTAYDVVDVDSHENPMGTILFGGTCGAPAPAAAGAGAGAGEEIPAPL 60
 QY 61 AGTVSKILVEGDTYKAGQTVLLEAMKMETEINAPTDGKVEKVLKERDAVGGGGLIK 120
 DB 61 AGTVSKILVEGDTYKAGQTVLLEAMKMETEINAPTDGKVEKVLKERDAVGGGGLIK 120
 QY 121 IG 122
 DB 121 IG 122

RESULT 6

AA28177
 ID AA28177 standard; Protein; ~~159~~-AA.

XX AA28177;

XX 18-MAR-1993 (first entry)

XX 1.3S polypeptide of Propionibacterium shermanii.

XX avidin-binding; biotin; protein purification;
 KW affinity chromatography.

XX Propionibacterium shermanii.

XX Key Location/Qualifiers
 FH 58.100
 FT /Label= biotin-binding_recognition_sequence
 XX
 PN

XX 04-NOV-1992.
 PD 07-APR-1992; 92EP-0303067
 XX 19-APR-1991; 91US-0687819.
 XX FROM J. ROHM & HAAS CO.
 XX CRESS DE, Haase FC;
 XX WPI: 1992-367575/45
 XX N-PSDB; AAQ29975.
 DR
 XX Hybrid polypeptide - contains a polypeptide fused to an avidin
 PT binding polypeptide contg. a biotin attachment domain
 PT
 PS Claim 6; Page 23; 40pp: English.

CC The P. shermanii 1.3S polypeptide is a preferred avidin-binding
 CC polypeptide for inclusion in the recombinant hybrid polypeptide of
 CC the invention. In the hybrid, a polypeptide of interest is fused to
 CC the C-terminus of the avidin-binding polypeptide. The hybrid
 CC polypeptide can then be recovered in a single chromatographic step
 CC using avidin monomer affinity chromatography.

SO Sequence 123 AA;

Query Match 98.3%; Score 604.5; DB 13; Length 123;

Best Local Similarity 99.2%; Pred. No. 2.4e-51; Mismatches 0; Indels 1; Gaps 1;

QY 1 MKLKVYNGTAYDVVDVDSHENPMGTILFGGTCGAPAPAAAGAGAGAGEEIPAP 59
 DB 1 MKLKVYNGTAYDVVDVDSHENPMGTILFGGTCGAPAPAAAGAGAGAGEEIPAP 60
 QY 60 LAGTVSKILVEGDTYKAGQTVLLEAMKMETEINAPTDGKVEKVLKERDAVGGGGLIK 119
 DB 61 LAGTVSKILVEGDTYKAGQTVLLEAMKMETEINAPTDGKVEKVLKERDAVGGGGLIK 120
 QY 120 KIG 122
 DB 121 KIG 123

RESULT 7

AA51122
 ID AA51122 standard; Protein; 125 AA.

XX AA51122;

XX 27-FEB-2002 (first entry)

XX Propionibacterium acnes immunogenic protein #12018.

XX SAPHO syndrome; synovitis; acne; pustulosis; hyperostosis; osteomyelitis;
 KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
 KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
 KW dermatological; osteopathic; neuroprotectant.

XX Propionibacterium acnes.

XX WO200181581-A2.

XX 01-NOV-2001.

XX 20-APR-2001; 2001WO-US12865.

XX 21-APR-2000; 2000US-199047P.

XX 02-JUN-2000; 2000US-208841P.
 XX 07-JUL-2000; 2000US-216747P.

PA (CORI-) CORIXA CORP.
 XX
 XX Skeiky YAM, Persing DH, Mitcham JL, Wang SS, Bhatia A;
 PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
 XX
 XX WPI: 2001-616774/71.
 DR N-PSDB: AAS59550.
 XX
 XX Propionibacterium acnes polypeptides and nucleic acids useful for
 PT vaccinating against and diagnosing infections, especially useful for
 PT treating acne vulgaris -
 XX
 XX Claim 6; SEQ ID NO 12317; 1069pp; English.
 XX
 XX Sequences AA039105-AA068017 represent Propionibacterium acnes immunogenic
 CC polypeptides. The proteins and their associated DNA sequences are used in
 CC the treatment, prevention and diagnosis of medical conditions caused by
 CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
 CC pustulosis, hyperostosis and osteomyelitis), uveitis and endophthalmitis.
 CC P. acnes is also involved in infections of bone, joints and the central
 CC nervous system, however it is particularly involved in the inflammatory
 CC lesions associated with acne vulgaris. A method for detecting the
 CC presence or absence of P. acnes in a patient comprises contacting a
 CC sample with a binding agent that binds to the proteins of the invention
 CC and determining the amount of bound protein in the sample. The
 CC polypeptides may be used as antigens in the production of antibodies
 CC specific for P. acnes proteins. These antibodies can be used to
 CC downregulate expression and activity of P. acnes polypeptides and
 CC therefore treat P. acnes infections. The antibodies may also be used as
 CC diagnostic agents for determining P. acnes presence, for example, by
 CC enzyme linked immunosorbent assay (ELISA).
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 XX Sequence 125 AA:
 SQ
 Query Match 70.9%; Score 436; DB 22; Length 125;
 Best Local Similarity 73.8%; Pred. No. 6.3e-35;
 Matches 90; Conservative 9; Mismatches 21; Indels 2; Gaps 2;
 QY 1 MKLKTAVNGTAVDVVDVDSKSHENMGITLFGGTGAPAPAPAGAGAGAGGEGEIPAPL 60
 Db 6 MKLKTAVNDVAVDVVDVDDKTNAPMAPILFGGAGG-PMRASGG-GAGKAGEGEVPAPL 63
 QY 61 AGTVSKIIIVKGGDFYKACQVLYLEAMKETEINAPTDGKYEKVLVYKERDAVQGGGLIK 120
 Db 64 AGTVAKIIVKGGDNVAKAGQVLTLEAMKETEINAPADGTVKGLIVANGDAVQGGGLVA 123
 QY 121 IG 122
 Db 124 LG 125
 RESULT 8
 AAB96610
 ID AAB96610 standard; Protein; 148 AA.
 XX
 XX AAB96610;
 AC
 XX
 XX 29-OCT-2001 (first entry)
 DT
 XX
 XX Putative biotin carboxylase carrier protein of acetyl-CoA carboxylase.
 DE
 XX
 XX Hyperthermophilic archaeon; hyperthermophilic protein.
 KW
 XX
 XX Pyrococcus abyssi.
 OS
 XX
 XX FR2792651-A1.
 PN
 XX
 XX 27-OCT-2000.
 PD
 XX
 XX 21-APR-1999; 99FR-0005034.
 PF

XX
 XX 21-APR-1999; 99FR-0005034.
 XX
 XX (CNRS) CNRS CENT NAT RECH SCI.
 PA (IFRE-) IFREMER INST FR RECH EXPL MER.
 XX
 XX Forterre P, Thierry JC, Prieur D, Dietrich J, Lecompte O;
 PI Querrelou J, Weissenbach J, Saurin W, Hellig R;
 DR WPI: 2001-126236/14.
 XX
 XX New nucleotide sequences isolated from Pyrococcus abyssi encode
 PT proteins useful in industry -
 PT
 XX
 XX Claim 7; Page 1349; 1657pp; French.
 XX
 XX The present invention relates to the genomic sequence of Pyrococcus
 CC abyssi (see AAF86431 and AAH4123-7) and P. abyssi proteins. P. abyssi is
 CC a hyperthermophilic archaeon, which is isolated from deep-sea
 CC hydrothermal vents. The present sequence is one such P. abyssi protein.
 CC The proteins of the present invention have various potential industrial
 CC uses, since the proteins are stable at very high temperatures, some up to
 CC 110 degrees centigrade.
 CC Note: This patent is in the same patent family as WO200065062, which
 CC contains additional sequences as shown in AAB99132-AAB99143,
 CC AAH75903-AAH75920 and AAG66436.
 XX
 XX Sequence 148 AA;
 SQ
 Query Match 37.6%; Score 231.5; DB 22; Length 148;
 Best Local Similarity 41.4%; Pred. No. 6.3e-15;
 Matches 60; Conservative 18; Mismatches 44; Indels 23; Gaps 3;
 QY 1 MKLKTAVNGTAVDVVDV-----VDSKSE-----NPMGTLFGGTGAPAP 40
 Db 4 MKKVVVNGKKEVDVVEEYMEKKEKRYTLEKTYEVAALGIVAPVQVATPAPTPTPT 63
 QY 41 PA---AGGAGAGKAGEGEIPAPLACTVSKIIIVKESDITVYKAGTVLYLEAMKETEINAP 97
 Db 64 PTPVQAPPTTPQVQASEENVVYAPMPGKVLKILVQEGQVYKLGCGLLILEAMKENEIPAPR 123
 QY 98 DGKYEKVLVYKERDAVQGGGLIKIG 122
 Db 124 DGVVKRILVKEGDVADVTGPTLEIG 148
 RESULT 9
 AAR42120
 ID AAR42120 standard; Protein; 100 AA.
 XX
 XX AAR42120;
 AC
 XX
 XX 05-MAY-1994 (first entry)
 DT
 XX
 XX Oxalacetate decarboxylase alpha subunit.
 DE
 XX
 XX Fusion protein; post-translation modification; biotination; marker.
 KW
 XX
 XX Klebsiella pneumoniae.
 OS
 XX
 XX US5252466-A.
 PN
 XX
 XX 12-OCT-1993.
 PD
 XX
 XX 18-MAY-1990; 90US-0525568.
 PF
 XX
 XX 19-MAY-1989; 89US-0354266.
 PR
 XX
 XX 18-MAY-1990; 90US-0525568.
 XX
 XX (BIOT-) BIOTECHNOLOGY RES & DEV CORP.
 PA (UNIV ILLINOIS FOUND.
 XX
 XX Cronan JE;
 PI

XX WPI; 1993-336075/42.
 XX
 XX Fusion proteins having site for post-translation modification -
 PT utilised esp. for biotin, used to identify or isolate fusion proteins
 PT from mixts.
 XX
 PS Claim 6; Page 57; 57pp: English.
 XX
 CC The sequence is that of a portion of the alpha subunit of Klebsiella
 CC pneumoniae oxalacetate decarboxylase. It is used as part of a fusion
 CC marker for the fusion protein that can be used directly or
 CC indirectly to identify the fusion protein or to isolate it from a
 CC mixt. of other materials such as host cell culture media.
 CC
 SQ Sequence 100 AA;
 Query Match 31.5%; Score 193.5; DB 14; Length 100;
 Best Local Similarity 51.8%; Pred. No. 1.9e-11;
 Matches 44; Conservative 10; Mismatches 30; Indels 1; Gaps 1;
 QY 38 AAPAPAGAGAGKAGGEG-ETIPAPLAGTIVSKILVEGDTVRAGOTVLYLEAMKMETEINAP 96
 DB 15 AAPAPASAPAAAAPAGACTPTAPLAGTIVKIVLASSEGOTVAAGEVLLILEAMKMETETIRAA 74
 QY 97 TDGKVEKVLKERNDAVGGGGLIKI 121
 DB 75 QAGTVRGIAVAKGDAVAVGDTLMTL 99
 RESULT 10
 AAB83169
 ID AAB83169 standard; Protein; 591 AA.
 XX
 AC AAB83169;
 XX
 DT 09-JUL-2001 (first entry)
 XX
 DE Corynebacterium thermoaminogenes accBC protein.
 XX
 KM Corynebacterium thermophilic; amino acid biosynthesis; enzyme;
 KM thermotolerant; aceA; accBC; dtsR1; dtsR2; ptk; scrB; glubCD;
 KM pdaA; pcc; ppc; acn; lcd; lpd; odhA.
 XX
 OS Corynebacterium thermoaminogenes.
 XX
 PN WO200125447-A1.
 XX
 PD 12-APR-2001.
 XX
 PE 04-OCT-2000; 2000WO-JP06913.
 XX
 PR 04-OCT-1999; 99JP-0282716.
 PR 01-NOV-1999; 99JP-0311147.
 PR 21-APR-2000; 2000JP-0120687.
 XX
 PA (AJIN) AJINOMOTO CO INC.
 XX
 PI Hirano S, Nonaka G, Matsuzaki Y, Akiyoshi N, Nakamura K, Kimura E;
 PI Osuni T, Matsui K, Kawahara Y, Kurahashi O, Nakamatsu T;
 PI Sugimoto S;
 DR WPI; 2001-300170/31.
 DR N-PSDB; AAF87428.
 XX
 PT Proteins and their DNA useful for microbial production of L-amino acids
 PT
 PS Claim 2; Page 91-93; 215pp; Japanese.
 XX
 CC The present sequence is provided in a specification relating to genes
 CC encoding thermophilic amino acid biosynthesis system enzymes of

CC the thermotolerant bacterium Corynebacterium thermoaminogenes.
 CC The novel proteins retain at least 30% isocitrate lyase activity
 CC after heating at 500C for 5 minutes. DNA fragments encoding the
 CC enzymes were isolated from a Corynebacterium thermoaminogenes
 CC chromosomal DNA plasmid library by PCR. The DNA may be used for
 CC developing strains of amino acid producing microorganisms.
 XX
 SQ Sequence 591 AA;
 Query Match 28.3%; Score 174; DB 22; Length 591;
 Best Local Similarity 36.3%; Pred. No. 1.4e-08;
 Matches 45; Conservative 20; Mismatches 45; Indels 14; Gaps 3;
 QY 2 KLVYVNGTAYDVVDVYKSHENPMGTILFSGGTGAPAP---AAGAGAGRAGEGEP 57
 DB 477 KVIYEIDGRREVALP-----GDIALGGAGAAKKPKRRRAGAKAGVSGD-SVA 526
 QY 58 APLAGTIVSKILVEGDTVRAGOTVLYLEAMKMETEINAPTDGKVEKVLKERNDAVGGG 117
 DB 527 APNQGTVIKVNVEDGAEVSEGDVTVLYLEAMKMPVAKHSGTVSGLTAAAGGVTKGOV 586
 QY 118 LIRI 121
 DB 587 LIRI 590
 RESULT 11
 AAU03126
 ID AAU03126 standard; Protein; 134 AA.
 XX
 AC AAU03126;
 XX
 DT 23-OCT-2001 (first entry)
 XX
 DE Streptococcus pyogenes virulence protein #8.
 XX
 KM Virulence; attenuated microorganism; Streptococcal infection;
 KM Gram-positive bacteria; antimicrobial; Impetigo; pneumonia.
 XX
 OS Streptococcus pyogenes.
 XX
 PN WO200148208-A2.
 PD 05-JUL-2001.
 XX
 PE 22-DEC-2000; 2000WO-GB04997.
 XX
 PF 23-DEC-1999; 99GB-0030462.
 PR 23-DEC-1999; 99GB-0030463.
 PR 23-DEC-1999; 99GB-0030464.
 PR 23-DEC-1999; 99GB-0030466.
 PR 23-DEC-1999; 99GB-0030467.
 PR 23-DEC-1999; 99GB-0030469.
 PR 23-DEC-1999; 99GB-0030471.
 PR 23-DEC-1999; 99GB-0030472.
 PR 23-DEC-1999; 99GB-0030473.
 PR 23-DEC-1999; 99GB-0030474.
 PR 23-DEC-1999; 99GB-0030475.
 PR 17-FEB-2000; 2000GB-0003725.
 PR 17-FEB-2000; 2000GB-0003726.
 PR 17-FEB-2000; 2000GB-0003727.
 PR 17-FEB-2000; 2000GB-0003728.
 PR 17-FEB-2000; 2000GB-0003729.
 PR 17-FEB-2000; 2000GB-0003730.
 PR 17-FEB-2000; 2000GB-0003731.
 PR 17-FEB-2000; 2000GB-0003732.
 PR 17-FEB-2000; 2000GB-0003733.
 PR 02-MAY-2000; 2000GB-0010585.
 PR 02-MAY-2000; 2000GB-0010587.
 XX
 PA (MICR-) MICROSCIENCE LTD.
 XX
 PI Clarke EE, Zhou L, Shea JE, Feldman RG, Holden DW;
 XX

QY 2 KLTATVNGTADVDYVDKSHEN-----PMGTILFGGSGTGAPAPAGAGAGAGAGE- 55
 Db 480 KVVVEIDGRREVRSLPDLALNSGGCDPVGVI-----RRKPKPRRGATGAASGDA 533
 QY 56 IPAPLAGVSKILVKEGTVKAGOTVLVLEAMKMETEINAPTDGKVEKVLKENDAYGG 115
 Db 534 VTAPMOGTIVKFAVEEGEVVAGDLVYVLEAMKMETEINAPTDGKVEKVLKENDAYGG 593
 QY 116 QGLIKI 121
 Db 594 TVLAELI 599
 RESULT 15
 ID AAU98050 standard; Protein: 1157 AA.
 AC AAU98050;
 DT 27-APR-2002 (first entry)
 DE Corynebacterium feedback-resistant pyruvate carboxylase enzyme.
 KM Feedback-resistant; pyruvate carboxylase; enzyme;
 KW aspartic acid feedback inhibition resistant.
 OS Corynebacterium glutamicum.
 FH Location/Qualifiers
 FT 1.18
 FT /note= "Specifically claimed in claim 18"
 FT 164.176
 FT /note= "Specifically claimed in claim 18"
 FT 193.205
 FT /note= "Specifically claimed in claim 18"
 FT 217.229
 FT /note= "Specifically claimed in claim 18"
 FT 238.250
 FT /note= "Specifically claimed in claim 18"
 FT 466.478
 FT /note= "Specifically claimed in claim 18"
 PN WO200231158-A2.
 PD 18-APR-2002.
 PF 12-OCT-2001; 2001WO-US31893.
 PR 13-OCT-2000; 2000US-239913P.
 PA (ARCH) ARCHER-DANIELS MIDLAND CO.
 PI Hanke PD;
 PI WPI: 2002-463267/49.
 DR N-PSDB: ABK52832.
 XX Novel mutated, feedback resistant pyruvate carboxylase enzyme
 PT polypeptide, useful for producing amino acids e.g. L-lysine,
 PT L-threonine, L-glycine, L-glutamic acid, L-proline and L-methionine and
 PT L-isoleucine
 XX Claim 17; Fig 1; 42pp; English.
 PS The present invention relates to a new mutated, feedback-resistant
 CC pyruvate carboxylase enzyme. The invention is useful for producing an
 CC amino acid (e.g. L-Lys, L-Thr, L-Met, L-Ile, L-Glu, L-Arg and L-Pro),
 CC by culturing a host cell in a suitable media and separating the amino
 CC acid from the medium. The vector of the invention is useful for
 CC replacement of a wild-type pyruvate carboxylase gene, with a feedback
 CC resistant pyruvate carboxylase gene, in a Corynebacterium glutamicum, by
 CC replacing a genomic copy of the wild-type pyruvate carboxylase gene with
 CC a selectable marker gene through homologous recombination to form a first

CC recombination strain, and replacing the selectable marker gene in the
 CC first recombinant strain, with feedback resistant pyruvate carboxylase
 CC gene through homologous recombination to form a second recombinant
 CC strain, where the homologous recombination in the above steps, occurs
 CC between the host cell and the vector. The feedback-resistant pyruvate
 CC carboxylase enzyme is resistant to feedback inhibition from aspartic
 CC acid. The present amino acid sequence represents the feedback-resistant
 CC pyruvate carboxylase enzyme of the invention.
 SO Sequence 1157 AA;
 Query Match 26.4%; Score 162.5; DB 23; Length 1157;
 Best Local Similarity 44.6%; Pred. No. 4.3e-07;
 Matches 33; Conservative 14; Mismatches 26; Indels 1; Gaps 1;
 QY 48 AGKAGEGIRPAPLAGVSKILVKEGTVKAGOTVLVLEAMKMETEINAPTDGKVEKVLK 107
 Db 1084 ADSSNKGHVAPFRGVT-VVVAEGDEVKADDAVILVLEAMKMETEINAPTDGKVEKVLK 1142
 QY 108 ERDAVGGGGLIKI 121
 Db 1143 AATKVEGGDLIVV 1156
 RESULT 16
 ID AAB83180 standard; Protein: 1139 AA.
 AC AAB83180;
 DT 09-JUL-2001 (first entry)
 DE Corynebacterium thermoaminogenes pc protein.
 KW Corynebacterium; thermophilic; amino acid biosynthesis; enzyme;
 KW thermotolerant; aceA; aceC; dtsK1; dtsK2; ptk; scrB; gluABCD;
 KW pdaA; pc; ppc; acn; lcd; lpd; odhA.
 OS Corynebacterium thermoaminogenes.
 PN WO200125447-A1.
 PD 12-APR-2001.
 PF 04-OCT-2000; 2000WO-JP06913.
 PR 04-OCT-1999; 99JP-0282716.
 PR 01-NOV-1999; 99JP-0311147.
 PR 21-APR-2000; 2000JP-0120687.
 PA (AJIN) AJINOMOTO CO INC.
 PI Hirano S, Nonaka G, Matsuzaki Y, Akiyoshi N, Nakamura K, Kimura E;
 PI Osumi T, Matsui K, Kawahara Y, Kurahashi O, Nakamatsu T;
 PI Sugimoto S;
 DR WPI: 2001-300170/31.
 DR N-PSDB: AAF87437.
 XX Proteins and their DNA useful for microbial production of L-amino acids
 PT
 PT
 PS Claim 9; Page 132-135; 215pp; Japanese.
 CC The present sequence is provided in a specification relating to genes
 CC encoding thermophilic amino acid biosynthesis system enzymes of
 CC the thermotolerant bacterium Corynebacterium thermoaminogenes.
 CC The novel proteins retain at least 30% isocitrate lyase activity
 CC after heating at 500C for 5 minutes. DNA fragments encoding the
 CC enzymes were isolated from a Corynebacterium thermoaminogenes
 CC chromosomal DNA plasmid library by PCR. The DNA may be used for
 CC developing strains of amino acid producing microorganisms.

PR 25-JUN-1999; 99US-0141031.
 PR 08-JUL-1999; 99DE-1031412.
 PR 08-JUL-1999; 99DE-1031413.
 PR 08-JUL-1999; 99DE-1031419.
 PR 08-JUL-1999; 99DE-1031420.
 PR 08-JUL-1999; 99DE-1031424.
 PR 08-JUL-1999; 99DE-1031428.
 PR 08-JUL-1999; 99DE-1031431.
 PR 08-JUL-1999; 99DE-1031433.
 PR 08-JUL-1999; 99DE-1031434.
 PR 08-JUL-1999; 99DE-1031510.
 PR 08-JUL-1999; 99DE-1031562.
 PR 08-JUL-1999; 99DE-1031634.
 PR 09-JUL-1999; 99DE-1032180.
 PR 09-JUL-1999; 99DE-1032227.
 PR 09-JUL-1999; 99DE-1032230.
 PR 09-JUL-1999; 99US-0143308.
 PR 14-JUL-1999; 99DE-1032824.
 PR 14-JUL-1999; 99DE-1032973.
 PR 14-JUL-1999; 99DE-1033005.
 PR 27-AUG-1999; 99DE-1040765.
 PR 31-AUG-1999; 99US-0151572.
 PR 03-SEP-1999; 99DE-1042076.
 PR 03-SEP-1999; 99DE-1042079.
 PR 03-SEP-1999; 99DE-1042086.
 PR 03-SEP-1999; 99DE-1042087.
 PR 03-SEP-1999; 99DE-1042088.
 PR 03-SEP-1999; 99DE-1042095.
 PR 03-SEP-1999; 99DE-1042123.
 PR 03-SEP-1999; 99DE-1042125.
 PA (BAD) BASF AG.
 XX
 XX
 PI Pompejus M, Kroeger B, Schroeder H, Zelder O, Haberhauer G;
 XX WPI; 2001-061975/07.
 DR N-PSDB; AAF71416.
 XX

PT New isolated Corynebacterium glutamicum nucleic acid encoding a sugar
 PT metabolism and oxidative phosphorylation protein for production or
 PT modulation of production of fine chemicals e.g. amino acids,
 PT carbohydrates or enzymes -
 XX
 XX
 PS Claim 20; Page 310-311; 1246pp; English.

CC AAF71360 to AAF71750 encode the Corynebacterium glutamicum sugar
 CC metabolism and oxidative phosphorylation (SMP) proteins given in
 CC AAB79243 to AAB 79633 which are involved in carbon metabolism and
 CC energy production. The C. glutamicum SMP gene can be used in vectors
 CC (II) for expression in host cells and production or modulation of
 CC production of fine chemicals, such as, an organic acid, a proteinogenic
 CC or nonproteinogenic amino acid (preferred), a purine or pyrimidine base,
 CC a nucleoside, a nucleotide, a lipid, a saturated or unsaturated fatty
 CC acid, a diol, a carbohydrate, an aromatic compound, a vitamin, a
 CC cofactor, a polyketide, or an enzyme. The presence of (I) or SMP proteins
 CC (III) encoded by them are used for diagnosing the presence or activity of
 CC Corynebacterium diphtheriae in a subject. (I), (II), (III) or host cells
 CC containing them are used to map genomes of organisms related to
 CC C. glutamicum, identity and localise C. glutamicum sequences of interest,
 CC in evolutionary studies, in determining SMP protein regions required
 CC for function, in modulating SMP protein activity, in modulating the
 CC metabolism of sugars, and in modulating high-energy molecule production
 CC in a cell (i.e. ATP, NADPH).
 CC
 XX
 SO Sequence 272 AA;

Query Match 25.9%; Score 159.5; DB 22; Length 272;
 Best Local Similarity 43.2%; Pred. No. 1.4e-07;
 Matches 32; Conservative 15; Mismatches 26; Indels 1; Gaps 1;

OY 48 AGKAGEEIPAPLAGTYSKILVKEGDTVKAGQVLTLEAMKMEETINAPDGGKVEKYLK 107
 DB 199 ADSSNKGHVAFAGVVT-VVAGDEVKAGDAVAITLAKMEETITASVDGKIDRVVVP 257

OY 108 ERDAVGGGGLIKI 121
 DB 258 AATFVEGDLIVV 271

RESULT 19
 ID AAW93971 standard; Protein; 1140 AA.
 XX AAW93971;
 AC AAW93971;
 XX
 DT 30-JUN-1999 (first entry)
 XX
 DE C. glutamicum pyruvate carboxylase protein.
 XX
 KW Pyruvate carboxylase; amino acid production; lysine production;
 KW threonine production; homoserine production; glutamate production;
 KW arginine production; feed additive; condiment; pharmaceutical;
 KW fine chemical; ss.
 XX
 OS Corynebacterium glutamicum.
 XX
 XX DE19831609-A1.
 PN
 XX
 PD 15-APR-1999.
 XX
 PE 14-JUL-1998; 98DE-1031609.
 XX
 PR 04-OCT-1997; 97DE-1043894.
 XX
 PA (KERJ) FORSCHUNGSZENTRUM JUELICH GMBH.
 XX
 PI Elkmanns B, Peters-Wendisch P, Sahm H;
 XX WPI; 1999-245521/21.
 DR N-PSDB; AAX24102.
 XX

PT Increasing microbial production of specific amino acids by
 PT increasing activity or expression of pyruvate carboxylase
 XX
 XX
 PS Claim 15; Page 11-15; 18pp; German.

CC This invention describes the isolation of a pyruvate carboxylase
 CC from Corynebacterium glutamicum which is used in a novel method for
 CC production of lysine, threonine, homoserine, glutamate and/or arginine,
 CC variously useful as feed additives, condiments, pharmaceuticals and
 CC intermediates for fine chemicals. Increasing pyruvate carboxylase
 CC activity increases the yield of microbial production of amino acids
 CC of the aspartate and/or glutamate families, e.g. about 50% more lysine,
 CC 40% more threonine and 150% more homoserine are secreted into the
 CC culture medium.
 CC
 XX
 SO Sequence 1140 AA;

Query Match 25.9%; Score 159.5; DB 20; Length 1140;
 Best Local Similarity 43.2%; Pred. No. 8.3e-07;
 Matches 32; Conservative 15; Mismatches 26; Indels 1; Gaps 1;

OY 48 AGKAGEEIPAPLAGTYSKILVKEGDTVKAGQVLTLEAMKMEETINAPDGGKVEKYLK 107
 DB 1067 ADSSNKGHVAFAGVVT-VVAGDEVKAGDAVAITLAKMEETITASVDGKIDRVVVP 1125
 OY 108 ERDAVGGGGLIKI 121
 DB 1126 AATFVEGDLIVV 1139

RESULT 20
 ID AAG90511 standard; Protein; 1140 AA.
 XX
 AC AAG90511;

```

XX 26-SEP-2001 (first entry)
DT C glutamicum protein fragment SEQ ID NO: 4265.
XX
DE Coryneform bacterium; amino acid synthesis; vitamin; saccharide;
XX organic acid synthesis.
XX
KM Corynebacterium glutamicum.
XX
OS EP1108790-A2.
XX
PN 20-JUN-2001.
XX
PD 18-DEC-2000; 2000EP-0127688.
XX
PF 16-DEC-1999; 99JP-0377484.
XX
PR 07-APR-2000; 2000JP-0159162.
XX
PR 03-AUG-2000; 2000JP-0280988.
XX
PA (KYOW ) KYOWA HAKKO KOGYO KK.
XX
PI Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
PI Tateishi N, Senoh A, Ikeda M, Ozaki A;
XX
DR WPI: 2001-376931/40.
XX
DR N-PSDB; AAH5730.
XX
PT Novel polynucleotides derived from Coryneform bacteria, for identifying
PT mutation point of a gene, measuring expression of a gene, analysing
PT expression profile or pattern of a gene and identifying homologous gene
XX
XX
PS Claim 17: SEQ ID NO: 4265; 246pp + Sequence Listing; English.
XX
XX
CC The present invention provides a number of nucleotide and protein
CC sequences from the Coryneform bacterium Corynebacterium glutamicum. These
CC are useful for identifying the mutation point of a gene derived from a
CC mutant of coryneform bacterium, measuring expression amount and
CC analysing the expression profile or expression pattern of a gene derived
CC from Coryneform bacterium, and identifying a homologue of a gene derived
CC from coryneform bacterium. Coryneform bacteria are useful for producing
CC amino acids, nucleic acids, vitamins, saccharides and organic acids,
CC particularly L-lysine. The present sequence is a protein described
CC in the exemplification of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from the
CC European Patent Office.
XX
SQ Sequence 1140 AA;

```

Query Match 25.9%; Score 159.5; DB 22; Length 1140;
 Best Local Similarity 43.2%; Pred. No. 8.3e-07;
 Matches 32; Conservative 15; Mismatches 26; Indels 1; Gaps 1;

```

OY 48 AGACGCEIPAPLAGTIVSKLVKESDTVAKGDTVYLVEAMKMETINAPDGVKVEVLK 107
   | : | : | | | | | : | | | | | | | | | | | | | | | | | | | | |
DB 1067 ADSNKGHVAAPFAGVYT-VTVAEGDEVRAGDAVAITTEAMKKEATTASVDKIDRVVVP 1125
OY 108 ERDAVGGGGLIKI 121
   | : | | : : :
DB 1126 AATKVEGGDLIVV 1139

```

RESULT 21
 AAG93249
 ID AAG93249 standard; Protein; 1140 AA.
 XX AAG93249;
 AC
 XX 26-SEP-2001 (first entry)
 DT
 XX C glutamicum protein fragment mutant P458S.
 DE

```

XX Coryneform bacterium; amino acid synthesis; vitamin; saccharide;
XX organic acid synthesis; mutant; mutein.
XX
KM Corynebacterium glutamicum.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 458
FT /note="Wild-type Pro substituted by Ser"
XX
XX
PN EP1108790-A2.
XX
PD 20-JUN-2001.
XX
PF 18-DEC-2000; 2000EP-0127688.
XX
PR 16-DEC-1999; 99JP-0377484.
XX
PR 07-APR-2000; 2000JP-0159162.
XX
PR 03-AUG-2000; 2000JP-0280988.
XX
PA (KYOW ) KYOWA HAKKO KOGYO KK.
XX
PI Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
PI Tateishi N, Senoh A, Ikeda M, Ozaki A;
XX
DR WPI: 2001-376931/40.
XX
DR N-PSDB; AAH5730.
XX
PT Novel polynucleotides derived from Coryneform bacteria, for identifying
PT mutation point of a gene, measuring expression of a gene, analysing
PT expression profile or pattern of a gene and identifying homologous gene
XX
XX
PS Claim 43: Page -: 246pp + Sequence Listing; English.
XX
XX
CC The present invention provides a number of nucleotide and protein
CC sequences from the Coryneform bacterium Corynebacterium glutamicum. These
CC are useful for identifying the mutation point of a gene derived from a
CC mutant of coryneform bacterium, measuring expression amount and
CC analysing the expression profile or expression pattern of a gene derived
CC from Coryneform bacterium, and identifying a homologue of a gene derived
CC from coryneform bacterium. Coryneform bacteria are useful for producing
CC amino acids, nucleic acids, vitamins, saccharides and organic acids,
CC particularly L-lysine. The present sequence is a mutant protein described
CC in the exemplification of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from the
CC European Patent Office.
XX
SQ Sequence 1140 AA;

```

Query Match 25.9%; Score 159.5; DB 22; Length 1140;
 Best Local Similarity 43.2%; Pred. No. 8.3e-07;
 Matches 32; Conservative 15; Mismatches 26; Indels 1; Gaps 1;

```

OY 48 AGACGCEIPAPLAGTIVSKLVKESDTVAKGDTVYLVEAMKMETINAPDGVKVEVLK 107
   | : | | | | | : | | | | | | | | | | | | | | | | | | | | |
DB 1067 ADSNKGHVAAPFAGVYT-VTVAEGDEVRAGDAVAITTEAMKKEATTASVDKIDRVVVP 1125
OY 108 ERDAVGGGGLIKI 121
   | : | | : : :
DB 1126 AATKVEGGDLIVV 1139

```

RESULT 22
 AAB67129
 ID AAB67129 standard; Protein; 1140 AA.
 XX AAB67129;
 AC
 XX 12-APR-2001 (first entry)
 DT
 XX Corynebacterium glutamicum pyruvate carboxylase.
 DE

DR WPI: 2000-465746/40.
DR N-PSDB; AAA47533.
XX Novel polynucleotides encoding Corynebacterium glutamicum pyruvate
PT carboxylase useful for industrial fermentation processes comprises a
PT specific nucleotide sequence
XX
PS Claim 3; Fig 1; 51pp; English.
XX
CC The pyruvate carboxylase of Corynebacterium glutamicum can be used
CC for producing amino acids, preferably lysine and glutamic acid in
CC industrial fermentations and for replenishing oxaloacetate consumed
CC for biosynthesis during growth. By incorporating the pyruvate
CC carboxylase gene in expression vectors levels of expression can be
CC 2 - 20 fold higher than in Corynebacterium glutamicum.
XX
SQ Sequence 1141 AA:

Query Match 25.9%; Score 159.5; DB 21; Length 1141;
Best Local Similarity 43.2%; Pred. No. 8.3e-07;
Matches 32; Conservative 15; Mismatches 26; Indels 1; Gaps 1;

QY 48 AGRAGEEITAPLAGTYSKILVKGDTYKAGQIVLLEAMKETEINAPTDGKYEKLVK 107
Db 1068 ADSSNKGHVAAPFAGVVT-VTVAGDEVKAGDAVAITEAMKMEATITASVDKIDRVVVP 1126
QY 108 ERDAVGGGGLIKI 121
Db 1127 AATKVEGGDLIVV 1140

RESULT 25
AA01086
ID AA01086 standard; protein; 161 AA.
XX
AC AA01086;
XX
DT 02-OCT-2001 (first entry)
XX
DE CFE 89 protein sequence.
XX
KW Antibacterial; vaccine; gene therapy; bacterial cell wall viability;
KW CFE; CFE; Conserved Essential Gene; bacterial infection;
KW antisense therapy; antibiotic resistance.
XX
OS Streptococcus pneumoniae.
XX
PN WO200149721-A2.
XX
PD 12-JUL-2001.
XX
PF 29-DEC-2000; 2000MO-US35604.
XX
PR 30-DEC-1999; 99US-0174089.
XX
PA (BRIM) BRISTOL-MYERS SQUIBB CO.
XX
PI Dougherty TJ, Pucci MJ, Dougherty BA, Davison DB, Bruccoleri RE;
PI Thannassi JA;
XX
DR WPI: 2001-496721/54.
DR N-PSDB; AAH90785.
XX
PT Nucleic acids encoding conserved essential genes involved in bacterial
PT replication which are potential targets for the treatment of antibiotic
PT resistant bacterial infections -
XX
PS Claim 27; Page 333; 380pp; English.
XX
CC The present invention relates to nucleic acids (AAH90701-AAH90918)
CC encoding polypeptides (AA01002-AA01114), which are essential for the
CC viability of a bacterial cell wall. The acronym CFE stands for "CEG For
CC Expression", where CEG stands for "Conserved Essential Gene". The nucleic

CC acids are useful for detecting the presence of proteins essential for the
CC viability of a bacterial cell wall in samples such as cells, tissues,
CC biological fluids, blood, serum, nose, ear or throat swabs with ligands,
CC and for detecting corresponding target nucleic acid molecules with
CC complementary sequences. The nucleic acids are also useful for
CC determining whether a genomic nucleotide sequence of interest is
CC essential for viability of a bacterial cell or whether it resides within
CC an operon, by integrating an exogenous nucleotide sequence comprising a
CC portion of an open reading frame of the genomic sequence of interest
CC (comprising 200-500 base pairs) into the genomic sequence of interest
CC which confers a selectable phenotype to the cell, and determining cell
CC viability with a selection agent such as chloramphenicol. The nucleic
CC acids and proteins are also useful as vaccines and for treating bacterial
CC infections with gene therapy and antisense therapy. The nucleic acids
CC also enable identification of targets suitable for the treatment of
CC antibiotic resistant bacterial infections.
XX
SQ Sequence 161 AA:

Query Match 25.9%; Score 159; DB 22; Length 161;
Best Local Similarity 41.8%; Pred. No. 8e-08;
Matches 38; Conservative 15; Mismatches 26; Indels 12; Gaps 3;

QY 39 PAPAGGAGAGKAGEGE-IPAPLAGTY-----SKLVKSGPTVYKAGQIVLLEAMKME 90
Db 74 PAPAE-----ASVASEGALVSPVGVYTLAAGPDKPAFTVGSVKQGTLLVTEAMKV 129
QY 91 TEINAPTDGKYEKLVNERDAVGGGGLIKI 121
Db 130 NEIPAPKIDGVVTELVSNEMVEFGKLVRI 160

RESULT 26
AA85871
ID AA85871 standard; protein; 163 AA.
XX
AC AA85871;
XX
DT 10-APR-2000 (first entry)
XX
DE S. pneumoniae derived protein #80.
XX
KW Treatment; prevention; disease; diagnosis; gene therapy; screening;
KW bacterial; antimicrobial; antibiotic; pathogenesis; infection.
XX
OS Streptococcus pneumoniae.
XX
PN WO9806734-A1.
XX
PD 19-FEB-1998.
XX
PF 15-AUG-1997; 97WO-US14436.
XX
PR 16-AUG-1996; 96US-0024022.
XX
PA (SMIK) SMITHKLINE BEECHAM CORP.
XX
PI Black MT, Hodgson JE, Knowles DUC, Lonetto MA, Nicholas RO;
PI Stodola RK;
XX
DR WPI: 1998-159452/14.
DR N-PSDB; AA96245.
XX
PT Streptococcus pneumoniae proteins and related DNA - useful for
PT screening compounds for antibacterial activity
XX
PS Claim 5; Page 374; 640pp; English.
XX
CC This invention describes novel isolated Streptococcus pneumoniae
CC polynucleotides (see AA96173-96494) and their encoded proteins (see
CC AA95792-986182). The DNA, vectors and host cells described in the
CC method of the invention are useful for the recombinant expression of the
CC polypeptides. The polypeptides are useful for treatment or prevention of

disease, or diagnosis of disease related to expression or activity of such a polypeptide. They can also be used to screen for compounds which interact with and inhibit or activate such a polypeptide. The polypeptides (or DNA encoding them, via gene therapy) are also useful for inducing an immunological response in a mammal. The antagonists are useful to inhibit such bacterial polypeptides. The polypeptides are particularly useful to identify antimicrobial compounds and antibiotics. They are also useful to determine their role in pathogenesis of infection, dysfunction and disease.

Sequence 163 AA:

Query Match 25.9%; Score 159; DB 19; Length 163;
Best Local Similarity 41.8%; Pred. No. 8.2e-08;
Matches 38; Conservative 15; Mismatches 26; Indels 12; Gaps 3;

OY 39 PAPAAGAGAGAGAGE-IPAPLAGTV-----SKLVKEDGYKAGOTVLEAMKE 90
DB 76 PAPAE---ASVASEGNLVESPLVGVYLAAGPDKPAPVTVGDSVKKQOTLVIITAMRV 131
OY 91 TEINAPTDGKYEKVLVKERDAVQGGGLIKI 121
DB 132 NEIPAPKDGVTTEILVSNEEMVEFGKGLVRI 162

RESULT 27
AAU37770
ID AAU37770 standard; Protein: 161 AA.

AC AAU37770;
XX
DT 14-FEB-2002 (first entry)
XX
DE Streptococcus pneumoniae cellular proliferation protein #199.
XX
KW Antisense; prokaryotic cellular proliferation protein;
KW antibiotic; antibacterial; drug design.
XX
OS Streptococcus pneumoniae.
XX
PN WO200170955-A2.
XX
PD 27-SEP-2001.
XX
PF 21-MAR-2001; 2001WO-US09180.
XX
PR 21-MAR-2000; 2000US-191078P.
PR 23-MAY-2000; 2000US-206848P.
PR 26-MAY-2000; 2000US-207727P.
PR 23-OCT-2000; 2000US-242578P.
PR 27-NOV-2000; 2000US-253625P.
PR 22-DEC-2000; 2000US-257931P.
PR 16-FEB-2001; 2001US-269308P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;
XX
DR WPI: 2001-611495/70.
DR N-PSDB: AAS55629.
XX
PT New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids -
XX
PS Example 3; Seq ID No 13363; 511pp; English.
XX
CC The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the
CC genes, their use in the discovery of novel antibiotics, the essential
CC genes themselves and the encoded proteins. The prokaryotes used are
CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The

Invention is also useful for the identification of potential new targets for antibiotic development. The antisense nucleic acids can also be used to identify proteins used in proliferation, to express these proteins, and to obtain antibodies capable of binding to the expressed proteins. The proteins can be used to screen compounds in rational drug discovery programmes. The antisense nucleic acid sequence is also useful to screen for homologous nucleic acids which are required for cell proliferation in a wide variety of organisms. The present sequence represents an essential prokaryotic cellular proliferation protein.
Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at [ftp.wipo.int/pub/published_pct_sequences](http://wipo.int/pub/published_pct_sequences).

Sequence 161 AA:

Query Match 25.5%; Score 157; DB 22; Length 161;
Best Local Similarity 41.8%; Pred. No. 1.3e-07;
Matches 38; Conservative 15; Mismatches 26; Indels 12; Gaps 3;

OY 39 PAPAAGAGAGAGAGE-IPAPLAGTV-----SKLVKEDGYKAGOTVLEAMKE 90
DB 74 PAPAE---ASVATEGNLVESPLVGVYLAAGPDKPAPVTVGDSVKKQOTLVIITAMRV 129
OY 91 TEINAPTDGKYEKVLVKERDAVQGGGLIKI 121
DB 130 NEIPAPKDGVTTEILVSNEEMVEFGKGLVRI 160

RESULT 28
AAU38022
ID AAU38022 standard; Protein: 161 AA.

AC AAU38022;
XX
DT 14-FEB-2002 (first entry)
XX
DE Streptococcus pneumoniae cellular proliferation protein #451.
XX
KW Antisense; prokaryotic cellular proliferation protein;
KW antibiotic; antibacterial; drug design.
XX
OS Streptococcus pneumoniae.
XX
PN WO200170955-A2.
XX
PD 27-SEP-2001.
XX
PF 21-MAR-2001; 2001WO-US09180.
XX
PR 21-MAR-2000; 2000US-191078P.
PR 23-MAY-2000; 2000US-206848P.
PR 26-MAY-2000; 2000US-207727P.
PR 23-OCT-2000; 2000US-242578P.
PR 27-NOV-2000; 2000US-253625P.
PR 22-DEC-2000; 2000US-257931P.
PR 16-FEB-2001; 2001US-269308P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;
XX
DR WPI: 2001-611495/70.
DR N-PSDB: AAS55881.
XX
PT New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids -
XX
PS Example 3; Seq ID No 13615; 511pp; English.
XX
CC The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the

PT New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids -

XX
PS Example 3; Seq ID No 12361; 511pp; English.

CC The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the
CC genes, their use in the discovery of novel antibiotics, the essential
CC genes themselves and the encoded proteins. The prokaryotes used are
CC *Escherichia coli*, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella*
CC *pneumoniae*, *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The
CC invention is also useful for the identification of potential new targets
CC for antibiotic development. The antisense nucleic acids can also be used
CC to identify proteins used in proliferation, to express these proteins,
CC and to obtain antibodies capable of binding to the expressed proteins.
CC The proteins can be used to screen compounds in rational drug discovery
CC programmes. The antisense nucleic acid sequence is also useful to screen
CC for homologous nucleic acids which are required for cell proliferation in
CC a wide variety of organisms. The present sequence represents an
CC essential prokaryotic cellular proliferation protein.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

XX
SQ Sequence 1073 AA;

Query Match 25.4%; Score 156; DB 22; Length 1073;

Best Local Similarity 29.2%; Pred. No. 1.7e-06;
Matches 42; Conservative 21; Mismatches 41; Indels 40; Gaps 3;

QY 8 NGTAYDVVD-----VDRSHENPMGTILFGGGTGAPAPAAAGAG----- 47
DB 935 NGETVEIIDKGRRLIKLEITISEPDENGNTIY-----AMNGARITIKDEN 984
QY 48 -----AGKAGEGEPAPLAGTYSKIIVKESDVIYKAGQIVLVEAMKETEINAPT 97
DB 985 VHTNANVKKPRADKSNPSHIGQMPGSVTEVYKAVKANQPLLTEAMMETTIOAPF 1044
QY 98 DGKYEKVLVKEKRDVAVGGGGLIKI 121
DB 1045 DGVIKQVTVNNGDTIATGDLLEI 1068

Search completed: May 1, 2003, 07:53:54
Job time : 43.3958 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4.D5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 1, 2003, 07:49:49 ; Search time 22.6042 Seconds
(without alignments)
412,647 Million cell updates/sec

Title: US-09-987-485-2

Perfect score: 342
Sequence: 1 EGRIPAPLAGVSKILVKEG.....KYLKRDVAGGQGLIKIG (70)

Scoring table: BLOSUM62
Gap 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 100 summaries

Database :

A_Geneseq_101002:.*
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	342	100.0	123	13	AA82177
2	342	100.0	126	21	AA84002
3	342	100.0	194	21	AA87898
4	342	100.0	253	18	AA81886
5	342	100.0	256	21	AA84001
6	342	100.0	266	21	AA84003
7	263	76.9	125	22	AA85122
8	195	57.0	148	22	AA89610
9	158.5	46.3	1157	23	AA898050
10	158	46.2	100	14	AA82120

11	157	45.9	1146	23	AB847612
12	155.5	45.5	272	22	AA87928
13	155.5	45.5	272	22	AA87928
14	155.5	45.5	1139	22	AA883180
15	155.5	45.5	1139	22	AA883180
16	155.5	45.5	1140	22	AA893971
17	155.5	45.5	1140	22	AA893971
18	155.5	45.5	1140	22	AA893971
19	155.5	45.5	1140	22	AA893971
20	155.5	45.5	1140	22	AA893971
21	151.5	44.3	1140	23	AA898052
22	151	44.2	166	23	AB228014
23	147	43.0	134	22	AA803126
24	146	42.7	1073	22	AA836768
25	146	42.7	1147	22	AA833972
26	144	42.1	108	20	AA873518
27	144	42.1	108	20	AA873518
28	144	42.1	161	22	AA837770
29	144	42.1	161	22	AA837770
30	144	42.1	161	22	AA837770
31	144	42.1	163	19	AA85871
32	144	42.1	243	20	AA814927
33	144	42.1	1142	22	AA835213
34	143	41.8	117	20	AA860030
35	143	41.8	294	22	AA823724
36	143	41.8	377	22	AA878644
37	143	41.8	725	22	AA894400
38	143	41.8	725	22	AA870157
39	143	41.8	725	22	AA870157
40	143	41.8	725	22	AA870157
41	142	41.5	155	22	AA833960
42	139.5	40.8	160	22	AA834089
43	139.5	40.8	162	22	AA834089
44	139	40.6	166	23	AA830189
45	139	40.6	184	23	AB288013
46	138	40.4	1154	22	AA82677
47	138	40.4	1154	22	AA82677
48	129	37.7	591	22	AA835853
49	129	37.7	1148	22	AA883169
50	127	37.1	600	21	AA800511
51	126.5	37.0	182	17	AA87874
52	126.5	37.0	182	17	AA87874
53	126	36.8	591	22	AA870404
54	126	36.8	591	22	AA870404
55	124.5	36.4	1181	22	AA876521
56	124.5	36.4	1181	22	AA876521
57	124.5	36.4	1181	22	AA876521
58	124.5	36.4	1181	22	AA876521
59	123	36.0	74	22	AA870404
60	122	35.7	164	22	AA870404
61	119.5	34.9	154	22	AA870404
62	119.5	34.9	154	22	AA870404
63	119.5	34.9	154	22	AA870404
64	119.5	34.9	154	22	AA870404
65	119.5	34.9	154	22	AA870404
66	118.5	34.6	101	23	AA874278
67	117.5	34.4	155	22	AA871776
68	116	33.9	590	23	AA898723
69	116	33.9	590	23	AA898723
70	116	33.9	1830	14	AA84504
71	115	33.6	71	19	AA872900
72	115	33.6	71	20	AA872900
73	113.5	33.2	630	22	AA872900
74	113.5	33.2	630	22	AA872900
75	111.5	32.6	111	23	AA872900
76	110	32.2	76	22	AA872900
77	110	32.2	76	22	AA872900
78	110	32.2	76	22	AA872900
79	110	32.2	76	22	AA872900
80	110	32.2	76	22	AA872900
81	110	32.2	76	22	AA872900
82	109	31.9	734	22	AA872900
83	109	31.9	22	16	AA872900

Listeria monocytogenes
Corynebacterium g1
Corynebacterium g1
Corynebacterium th
C. glutamicum pyru
C. glutamicum prote
C. glutamicum prote
Corynebacterium g1
Corynebacterium w1
pyruvate carboxyla
Corynebacterium mu
Streptococcus poly
Streptococcus pyog
Staphylococcus aur
Staphylococcus aur
M. vaccae antigen
M. vaccae GV-29 par
Streptococcus pneu
Streptococcus pneu
CFE 89 protein seq
S. pneumoniae derl
Amino acid sequenc
Enterococcus faeca
Human endometrium
Human endometrium
Novel human enzyme
CoA carboxylase 41
Human protein seg
DNA encoding human
Human biotin enzym
Lactococcus lactis
Lactococcus lactis
Enterococcus faeca
Enterococcus faeca
Streptococcus poly
Streptococcus poly
S. epidermidis ope
Staphylococcus epi
Corynebacterium th
Bacillus subtilis
M. tuberculosis an
Anabaena biotin ca
Anabaena biotin ca
C. glutamicum prote
Corynebacterium g1
Drosophila melanog
Drosophila melanog
Drosophila melanog
Drosophila melanog
Novel human secret
Protein involved i
Staphylococcus aur
Staphylococcus aur
Staphylococcus epi
Synecchococcus blot
Synecchococcus blot
Human biotin-depen
S. epidermidis ope
S. coelicolor Acca
S. coelicolor Acca
Urea amide lyase.
Mycobacterium tube
Amino acid sequenc
E. coli cellular p
Novel mar regulat
Human carboxylase
Human colon cancer
Human ovarian anti
Micromonospora eve
Human human diagno
Human polypeptide
Human polypeptide
Biotinylation pept
Propionibacterium

64 108.5 31.7 167 20 AAY34780
 65 108 31.6 146 22 AAG81854
 86 108 31.6 146 23 ABP38749
 87 104.5 30.6 155 23 ABP48312
 88 104 30.4 362 21 AAG17822
 89 104 30.4 364 21 AAG46637
 90 104 30.4 461 21 AAG17821
 91 104 30.4 462 21 AAG17820
 92 104 30.4 463 21 AAG46636
 93 104 30.4 464 21 AAG46635
 94 104 30.4 1835 14 AAR42839
 95 103 30.1 613 22 AAG62343
 96 102.5 30.0 153 22 AAG35366
 97 101.5 29.7 89 20 AAW90002
 98 101.5 29.7 156 20 AAW90001
 99 101.5 29.7 156 22 AAW35729
 100 101.5 29.7 160 20 AAW89895

ALIGNMENTS

Chlamydia pneumoniae
 S. epidermidis ope
 Staphylococcus epi
 Listeria monocytog
 Arabidopsis thaila
 Arabidopsis thaila
 Arabidopsis thaila
 Arabidopsis thaila
 Arabidopsis thaila
 Arabidopsis thaila
 Urea amidolyase
 Propionibacterium
 Haemophilus Influe
 Expressed antigen
 Expressed antigen
 Helicobacter pylori
 Antigen 1 from clu

RESULT 1

AA828177 standard; Protein: 123 AA.

AA828177;

18-MAR-1993 (first entry)

1.35 polypeptide of Propionibacterium shermanii.

avidin-binding; biotin; protein purification;
 affinity chromatography.

Propionibacterium shermanii.

Location/Qualifiers

Region 58..100 /label= biotin-binding_recognition_sequence

PN: BP511747-A.

04-NOV-1992.

07-APR-1992; 92EP-0303067.

19-APR-1991; 91US-0687819.

(ROHM-ROHM-4-HAAS-CC)

(KRS/) CRESS D E.

Cress DE, Haase FC;

WPI: 1992-367575/45.

N-PSDB: AAO29975.

Hybrid polypeptide - contains a polypeptide fused to an avidin
 binding polypeptide contg. a biotin attachment domain

Claim 6; Page 23; 40pp; English.

The P. shermanii 1.35 polypeptide is a preferred avidin-binding
 polypeptide for inclusion in the recombinant hybrid polypeptide of
 the invention. In the hybrid, a polypeptide of interest is fused to
 the C-terminus of the avidin-binding polypeptide. The hybrid
 polypeptide can then be recovered in a single chromatographic step
 using avidin monomer affinity chromatography.

Sequence 123 AA;

Query Match 100.0%; Score 342; DB 13; Length 123;
 Best Local Similarity 100.0%; Pred. No. 5.4e-34;

Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EGRIPLAGTYSKILVKGDTYKAGQTVLVEAMKMETEINAPTDGKVEKLVKRDV 60
 |||||
 Db 54 EGRIPLAGTYSKILVKGDTYKAGQTVLVEAMKMETEINAPTDGKVEKLVKRDV 113

QY 61 QGGGGLIKIG 70
 |||||
 Db 114 QGGGGLIKIG 123

RESULT 2

AA84002 standard; Protein: 126 AA.

AA84002;

03-JUL-2000 (first entry)

Amino acid sequence of a T. pallidum protein.

Pinpoint vector; fusion protein antigen; membrane antigen; syphilis.

Treponema pallidum.

PN: EP985931-A2

15-MAR-2000.

12-AUG-1999; 99EP-0115877.

04-SEP-1998; 98US-0148920.

(BECT) BECTON DICKINSON & CO.

Mullenix MC, Deutsch J;

WPI: 2000-226057/20.

N-PSDB: AA29247.

Immunosay for anti-Treponema pallidum antibodies, used for diagnosis
 of syphilis using a fusion protein of membrane antigen with peptide
 sequence that can be biotinylated in vivo

Disclosure; Page 12-13; 16pp; English.

The present sequence represents a Treponema pallidum protein, which is
 used to produce a fusion protein antigen for use in the method of
 the invention. The specification describes a method for detecting
 CC antibodies against Treponema pallidum. The antibodies are detected
 in a sample by reaction with a fusion protein antigen, present in
 the mixture in limiting concentration. The fusion protein antigen
 comprises a Treponema pallidum membrane antigen. The method is used
 for diagnosis of syphilis.

Sequence 126 AA;

Query Match 100.0%; Score 342; DB 21; Length 126;
 Best Local Similarity 100.0%; Pred. No. 5.5e-33;
 Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EGRIPLAGTYSKILVKGDTYKAGQTVLVEAMKMETEINAPTDGKVEKLVKRDV 60
 |||||
 Db 53 EGRIPLAGTYSKILVKGDTYKAGQTVLVEAMKMETEINAPTDGKVEKLVKRDV 112

QY 61 QGGGGLIKIG 70
 |||||
 Db 113 QGGGGLIKIG 122

RESULT 3

AA878908 standard; Protein: 194 AA.

XX AC AAY78908;
 XX DT 19-MAY-2000 (first entry)
 XX DE Outer surface protein C (OspC) DraI fragment amino acid sequence.
 XX KW Outer surface protein C; OspC; Immunological epitope: Lyme disease;
 XX KW vaccine; prevention; Borrelia infection; diagnose.
 XX OS Borrelia burgdorferi.
 XX PN W0200006745-A1.
 XX PD 10-FEB-2000.
 XX PF 30-JUL-1999; 99WO-US17270.
 XX PR 31-JUL-1998; 98US-0094955.
 XX PA (GUND-) GUNDERSEN LUTHERAN MEDICAL FOUND. INC.
 XX PI Callister SN, Lovrich SD, Schell RF, Jobe DA;
 XX DR WPI; 2000-195305/17.
 XX DR N-PSDB; AA292216.
 XX PT New immunogenic polypeptides useful as a vaccine against Lyme disease
 XX PT and for treating and detecting borrelia infection in mammals consists
 XX PT an epitope of Borrelia burgdorferi OspC fragment
 XX PS Claim 3; Fig 4; 51pp; English.
 XX CC This sequence represents the Borrelia burgdorferi outer surface protein C
 XX CC (OspC) DraI fragment amino acid sequence. The polypeptide contains an
 XX CC immunological epitope used in the invention. Large amounts of OspC are
 XX CC rapidly synthesized by B. burgdorferi shortly after attachment of
 XX CC infected ticks to mammalian hosts. The OspC protein sequence is used to
 XX CC diagnose B. burgdorferi infection in mammals. The OspC nucleotide sequence
 XX CC is used to prevent (via vaccination), treat or detect Borrelia
 XX CC (especially B. burgdorferi) infections, i.e. Lyme disease, in mammals
 XX CC including humans. The OspC nucleotide sequence provides a superior
 XX CC diagnostic antigen that detects early Lyme disease infection, predicts
 XX CC successful eradication or the organism from the host, and discriminates
 XX CC between individuals with Lyme disease and individuals who have been
 XX CC vaccinated with an OspA Lyme disease vaccination. Detection of anti-OspC
 XX CC borreliaclad antibodies advantageously gives an early diagnosis which
 XX CC anti-OspA and anti-OspB borreliaclad antibodies cannot do. Unlike
 XX CC vaccination with OspA, vaccination with OspC results in clearance of
 XX CC spirochetes and resolution of symptoms even if administered after
 XX CC infection with B. burgdorferi.
 XX SQ Sequence 194 AA;
 XX Query Match 100.0%; Score 342; DB 21; Length 194;
 XX Best Local Similarity 100.0%; Pred. No. 9.7e-34;
 XX Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 EGEPAPLAGVSKILVKEGDTVAKGOTVYLKMKMETETINAPTDGKVEKVLKENDAV 60
 DB 53 EGEPAPLAGVSKILVKEGDTVAKGOTVYLKMKMETETINAPTDGKVEKVLKENDAV 112
 OY 61 QGGGGLIKIG 70
 DB 113 QGGGGLIKIG 122

DT 21-APR-1997 (first entry)
 XX DE Hantavirus nuclear protein sequence.
 XX KW Hantavirus nuclear protein; antigen; monoclonal antibody;
 XX KW diagnosis; specificity.
 XX OS Hantavirus.
 XX PN JP08325291-A.
 XX PD 10-DEC-1996.
 XX PF 30-MAY-1995; 95JP-0132460.
 XX PR 30-MAY-1995; 95JP-0132460.
 XX PA (ATAT-) A & T KK.
 XX PA (ARIK/) ARIKAWA J.
 XX PA (HASH/) HASHIMOTO N.
 XX DR WPI; 1997-083468/08.
 XX PT Hantavirus antigen protein and monoclonal antibody - used in the
 XX PT diagnosis and treatment of hantavirus infection
 XX PS Example 1; Page 17; 18pp; Japanese.
 XX CC The sequences given in AAM11872-86 represent Hantavirus nuclear
 XX CC proteins. The N-terminal regions of these proteins may be used as
 XX CC antigens to raise anti-hantavirus monoclonal antibodies. These
 XX CC antibodies are useful in diagnosis of hantavirus infection due to
 XX CC their specificity to Hantavirus.
 XX SQ Sequence 253 AA;
 XX Query Match 100.0%; Score 342; DB 18; Length 253;
 XX Best Local Similarity 100.0%; Pred. No. 1.4e-33;
 XX Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 EGEPAPLAGVSKILVKEGDTVAKGOTVYLKMKMETETINAPTDGKVEKVLKENDAV 60
 DB 53 EGEPAPLAGVSKILVKEGDTVAKGOTVYLKMKMETETINAPTDGKVEKVLKENDAV 112
 OY 61 QGGGGLIKIG 70
 DB 113 QGGGGLIKIG 122

RESULT 5
 ID AAY84001 standard; Protein; 256 AA.
 AC AAY84001;
 DT 03-JUL-2000 (first entry)
 XX Amino acid sequence of a T. pallidum fusion protein antigen.
 DE Pinpoint vector; fusion protein antigen; membrane antigen; syphilis.
 XX OS Synthetic.
 XX OS Treponema pallidum.
 XX PN EP985931-A2.
 XX PD 15-MAR-2000.
 XX PF 12-AUG-1999; 99EP-0115877.
 XX PR 04-SEP-1998; 98US-0148920.
 XX PA (BECT) BECTON DICKINSON & CO.

XX Mullerix MC, Deutsch J;
 XX WPI: 2000-226057/20.
 DR N-PSDB; AA29246.
 XX
 PT Immunassay for anti-Treponema pallidum antibodies, used for diagnosis
 PT of syphilis using a fusion protein of membrane antigen with peptide
 PT sequence that can be biotinylated in vivo
 XX
 PS Claim 9; Page 10-11; 16pp; English.
 XX
 CC The present sequence represents a Treponema pallidum fusion protein
 CC antigen. The protein is used in the method of the invention. The
 CC specification describes a method for detecting antibodies against
 CC Treponema pallidum. The antibodies are detected in a sample by
 CC reaction with a fusion protein antigen, present in the mixture in
 CC limiting concentration. The fusion protein antigen comprises a
 CC Treponema pallidum membrane antigen. The method is used for diagnosis
 CC of syphilis.
 CC
 SQ Sequence 256 AA;

Query Match 100.0%; Score 342; DB 21; Length 256;
 Best Local Similarity 100.0%; Pred. No. 1.4e-33;
 Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EGEIPAPLAGTWSKILVKEGDFVAKAGTVLVEAMKMETEINAPTDGKVEKVLKERDAV 60
 DB 53 EGEIPAPLAGTWSKILVKEGDFVAKAGTVLVEAMKMETEINAPTDGKVEKVLKERDAV 112
 OY 61 QGGGGLIKIG 70
 DB 113 QGGGGLIKIG 122

RESULT 6

AAV84003
 ID AAV84003 standard; Protein; 266 AA.

XX AAV84003;
 AC
 XX
 DT 03-JUL-2000 (first entry)
 XX
 DE Amino acid sequence of a T. pallidum membrane protein antigen.
 XX
 KW Pinpoint vector; fusion protein antigen; membrane antigen; syphilis.
 XX
 OS Treponema pallidum.
 XX
 PN EP985931-A2.
 PD 15-MAR-2000.
 XX
 PF 12-AUG-1999; 98EP-0115877.
 XX
 PR 04-SEP-1998; 98OS-0148920.
 XX
 PA (BECT) BECTON DICKINSON & CO.
 XX
 PI Mullerix MC, Deutsch J;
 XX
 DR WPI: 2000-226057/20.
 DR N-PSDB; AA29248.
 XX
 PT Immunassay for anti-Treponema pallidum antibodies, used for diagnosis
 PT of syphilis using a fusion protein of membrane antigen with peptide
 PT sequence that can be biotinylated in vivo
 XX
 PS Claim 9; Page 14-15; 16pp; English.
 XX
 CC The present sequence represents a Treponema pallidum membrane protein
 CC antigen, which is used to produce a fusion protein antigen for use

CC In the method of the invention. The specification describes a method
 CC for detecting antibodies against Treponema pallidum. The antibodies
 CC are detected in a sample by reaction with a fusion protein antigen,
 CC present in the mixture in limiting concentration. The fusion protein
 CC antigen comprises a Treponema pallidum membrane antigen. The method is
 CC used for diagnosis of syphilis.
 CC
 SQ Sequence 266 AA;

Query Match 100.0%; Score 342; DB 21; Length 266;
 Best Local Similarity 100.0%; Pred. No. 1.5e-33;
 Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EGEIPAPLAGTWSKILVKEGDFVAKAGTVLVEAMKMETEINAPTDGKVEKVLKERDAV 60
 DB 53 EGEIPAPLAGTWSKILVKEGDFVAKAGTVLVEAMKMETEINAPTDGKVEKVLKERDAV 112
 OY 61 QGGGGLIKIG 70
 DB 113 QGGGGLIKIG 122

RESULT 7

AAU51122
 ID AAU51122 standard; Protein; 125 AA.

XX AAU51122;
 AC
 XX
 DT 27-FEB-2002 (first entry)
 XX
 DE Propionibacterium acnes immunogenic protein #12018.
 XX
 KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
 KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
 KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
 KW dermatological; osteopathic; neuroprotectant.
 XX
 OS Propionibacterium acnes.
 XX
 PN WO200181581-A2.
 PD 01-NOV-2001.
 XX
 PF 20-APR-2001; 2001WO-US12865.
 XX
 PR 21-APR-2000; 2000US-199047P.
 PR 02-JUN-2000; 2000US-208841P.
 PR 07-JUL-2000; 2000US-216747P.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Skelky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
 PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
 DR WPI: 2001-616774/71.
 DR N-PSDB; AAS59550.
 XX
 PT Propionibacterium acnes polypeptides and nucleic acids useful for
 PT vaccinating against and diagnosing infections, especially useful for
 PT treating acne vulgaris -
 XX
 PS Claim 6; SEQ ID No 12317; 1069pp; English.
 XX
 CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
 CC polypeptides. The proteins and their associated DNA sequences are used in
 CC the treatment, prevention and diagnosis of medical conditions caused by
 CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
 CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
 CC P. acnes is also involved in infections of bone, joints and the central
 CC nervous system, however it is particularly involved in the inflammatory
 CC lesions associated with acne vulgaris. A method for detecting the
 CC presence or absence of P. acnes in a patient comprises contacting a
 CC sample with a binding agent that binds to the proteins of the invention

CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for *P. acnes* proteins. These antibodies can be used to
CC downregulate expression and activity of *P. acnes* polypeptides and
CC therefore treat *P. acnes* infections. The antibodies may also be used as
CC diagnostic agents for determining *P. acnes* presence, for example, by
CC enzyme linked immunosorbent assay (ELISA).
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

SO Sequence 125 AA;

Query Match 76.9%; Score 263; DB 22; Length 125;
Best Local Similarity 74.3%; Pred. No. 2,4e-24;
Matches 52; Conservative 7; Mismatches 11; Indels 0; Gaps 0;

QY 1 EGEIPAPLAGVSKIIKEDGTVAAGTIVLLEAMKMETETINAPTDKVEVLYKEDAV 60
DB 56 EGEVPAPLAGVAKILVAGSDAVAGVLLTLEAMKMETETINAPDCTVKGILVAGDAV 115

OY 61 QGGGGLIKIG 70
DB 116 QGGGGLVATG 125

RESULT 8

AAB96610 standard; Protein; 148 AA.

AAB96610;

29-OCT-2001 (first entry)

Putative biotin carboxylase carrier protein of acetyl-CoA carboxylase.

Hyperthermophilic archaeon; hyperthermophilic protein.

Pyrococcus abyssi.

FR2792651-A1.

27-OCT-2000.

21-APR-1999; 99FR-0005034.

21-APR-1999; 99FR-0005034.

(CNRS) CNRS CENT NAT RECH SCI.

(IFRE-) IFREMER INST FR RECH EXPL MER.

Porterre P, Thierry JC, Prieur D, Dietrich J, Lecompte O;

Querellou J, Weissenbach J, Saurin W, Hellig R;

WPI; 2001-126236/14.

New nucleotide sequences isolated from *Pyrococcus abyssi* encode

proteins useful in industry -

Claim 7; Page 1349; 1657pp; French.

The present invention relates to the genomic sequence of *Pyrococcus*

abyssi (see AAB96431 and AAB1223-7) and *P. abyssi* proteins. *P. abyssi* is

a hyperthermophilic archaeon, which is isolated from deep-sea

hydrothermal vents. The present sequence is one such *P. abyssi* protein.

The proteins of the present invention have various potential industrial

uses, since the proteins are stable at very high temperatures, some up to

110 degrees centigrade.

Note: This patent is in the same patent family as WO200065062, which

contains additional sequences as shown in AAB99132-AAB99143,

AAB75903-AAB75920 and AAB66436.

Sequence 148 AA;

Query Match 57.0%; Score 195; DB 22; Length 148;
Best Local Similarity 58.6%; Pred. No. 6.1e-16;
Matches 41; Conservative 10; Mismatches 19; Indels 0; Gaps 0;

QY 1 EGEIPAPLAGVSKIIKEDGTVAAGTIVLLEAMKMETETINAPTDKVEVLYKEDAV 60
DB 79 ENVTAPMPGVKILVQGGVAVLGGGLILEAMKMETETINAPDGVVVKILVKGDAV 138

OY 61 QGGGGLIKIG 70
DB 139 DTGTPLEIG 148

RESULT 9

AAU98050 standard; Protein; 1157 AA.

AAU98050;

27-AUG-2002 (first entry)

Corynebacterium feedback-resistant pyruvate carboxylase enzyme.

Feedback-resistant; pyruvate carboxylase; enzyme;

aspartic acid feedback inhibition resistant.

Corynebacterium glutamicum.

Location/Qualifiers

1..18 /note- "Specifically claimed in claim 18"

164..176 /note- "Specifically claimed in claim 18"

193..205 /note- "Specifically claimed in claim 18"

217..229 /note- "Specifically claimed in claim 18"

238..250 /note- "Specifically claimed in claim 18"

466..478 /note- "Specifically claimed in claim 18"

WO200231158-A2.

18-APR-2002.

12-OCT-2001; 2001WO-US31893.

13-OCT-2000; 2000US-239913P.

(ARCH) ARCHER-DANIELS MIDLAND CO.

Hanke PD;

WPI; 2002-463267/49.

N-PSDB; ABK52832.

Novel mutated, feedback resistant pyruvate carboxylase enzyme

polypeptide, useful for producing amino acids e.g. L-lysine,

L-threonine, L-glutamine, L-glutamic acid, L-proline and L-methionine and

L-isoleucine -

Claim 17; Fig 1; 42pp; English.

The present invention relates to a new mutated, feedback-resistant

pyruvate carboxylase enzyme. The invention is useful for producing an

amino acid (e.g. L-Lys, L-Thr, L-Met, L-Ile, L-Glu, L-Arg and L-Pro),

by culturing a host cell in a suitable media and separating the amino

acid from the medium. The vector of the invention is useful for

replacement of a wild-type pyruvate carboxylase gene, with a feedback

resistant pyruvate carboxylase gene, in a *Corynebacterium glutamicum*, by

replacing a genomic copy of the wild-type pyruvate carboxylase gene with

CC a selectable marker gene through homologous recombination to form a first
 CC recombinant strain, and replacing the selectable marker gene in the
 CC first recombinant strain, with feedback resistant pyruvate carboxylase
 CC gene through homologous recombination to form a second recombinant
 CC strain, where the homologous recombination in the above steps, occurs
 CC between the host cell and the vector. The feedback-resistant pyruvate
 CC carboxylase enzyme is resistant to feedback inhibition from aspartic
 CC acid. The present amino acid sequence represents the feedback-resistant
 CC pyruvate carboxylase enzyme of the invention.

SQ Sequence 1157 AA:

Query Match 46.3%; Score 158.5; DB 23; Length 1157;
 Best Local Similarity 46.4%; Pred. No. 2.6e-10;
 Matches 32; Conservative 13; Mismatches 23; Indels 1; Gaps 1;

OY 1 EGSEIPPLACTGSKILVKGEDTVKAGOTVLTLEAMKMETEINAPTDGKYEKVLKERDAVGG 60
 DB 1089 KGHVAPFAGVVT-VTVAGDEYKAGDAVAITLAKMEATITTSVDSKIERVVPAATKV 1147
 OY 61 QGGGGLIKI 69
 DB 1148 EGGDLIVVV 1156

RESULT 10
 AAR42120

ID AAR42120 standard; protein: 100 AA.

AC AAR42120;

DT 05-MAY-1994 (first entry)

DE Oxalacetate decarboxylase alpha subunit.

KM Fusion protein; post-translation modification; biotination; marker.

OS Klebsiella pneumoniae.

PN US5252466-A.

PD 12-OCT-1993.

PF 18-MAY-1990; 90US-0525568.

PR 19-MAY-1989; 89US-0354266.

PR 18-MAY-1990; 90US-0525568.

PA (BIOT-) BIOTECHNOLOGY RES & DEV CORP.

PA (UNIT) UNIV ILLINOIS FOUND.

PI Cronan JE;

PI Cronan JE;

PI Cronan JE;

PI Cronan JE;

PI Cronan JE;

PI Cronan JE;

PI Cronan JE;

PI Cronan JE;

PI Cronan JE;

PI Cronan JE;

PI Cronan JE;

PI Cronan JE;

PI Cronan JE;

PI Cronan JE;

OY 4 IPAPLACTGSKILVKGEDTVKAGOTVLTLEAMKMETEINAPTDGKYEKVLKERDAVGG 63
 DB 34 VTPAPLACTGSKILVKGEDTVKAGOTVLTLEAMKMETEINAPTDGKYEKVLKERDAVGG 93
 OY 64 QGLIKI 69
 DB 94 DTLMTL 99

RESULT 11

ID ABB47612 standard; protein: 1146 AA.

AC ABB47612;

DT 05-FEB-2002 (first entry)

DE Listeria monocytogenes protein #316.

KM Antibacterial; gene therapy; vaccine; biosynthesis; biodegradation;

KM Vitamin B12; bacterial infection; disease.

OS Listeria monocytogenes.

PN WO200177335-A2.

PD 18-OCT-2001.

PF 11-APR-2001; 2001WO-FR01118.

PR 11-APR-2000; 2000FR-0004629.

PA (INSP) INST PASTEUR.

PA (INSP) INST PASTEUR.

PA (INSP) INST PASTEUR.

PA (INSP) INST PASTEUR.

PA (INSP) INST PASTEUR.

PA (INSP) INST PASTEUR.

PA (INSP) INST PASTEUR.

PA (INSP) INST PASTEUR.

PA (INSP) INST PASTEUR.

PA (INSP) INST PASTEUR.

PA (INSP) INST PASTEUR.

PA (INSP) INST PASTEUR.

PA (INSP) INST PASTEUR.

PA (INSP) INST PASTEUR.

PA (INSP) INST PASTEUR.

PA (INSP) INST PASTEUR.

PA (INSP) INST PASTEUR.

PA (INSP) INST PASTEUR.

PA (INSP) INST PASTEUR.

PA (INSP) INST PASTEUR.

PA (INSP) INST PASTEUR.

PA (INSP) INST PASTEUR.

PA (INSP) INST PASTEUR.

PA (INSP) INST PASTEUR.

Claim 6: SEQ ID NO 317; 192pp; French.

The present invention relates to the genome sequence of *Listeria monocytogenes* EGD-e (see ABR03041). The genome sequence and fragments of it are useful for selecting probes and primers for detecting genes in *L. monocytogenes* and related organisms, and for studying genetic polymorphisms and other genomes. The present sequence is a protein encoded by the genome sequence of the present invention. Proteins expressed from the genome sequence are useful for raising specific antibodies, identification of *L. monocytogenes* and related organisms, and for biosynthesis and biodegradation, especially biosynthesis of Vitamin B12. The genome sequence and proteins encoded by it are also useful for selecting compounds that regulate gene expression and cell replication and modulate *L. monocytogenes*-related diseases. In addition, the genome sequence and proteins encoded by it are useful in pharmaceutical and vaccine compositions for the treatment or prevention of infections by *L. monocytogenes* and related organisms.

Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

Sequence 1146 AA:

[illegible]

PS Claim 1; Page -: 42pp; English.

XX The present invention relates to a new mutated, feedback-resistant

CC pyruvate carboxylase enzyme. The invention is useful for producing an

CC amino acid (e.g. L-Lys, L-Thr, L-Met, L-Ile, L-Glu, L-Arg and L-Pro),

CC by culturing a host cell in a suitable media and separating the amino

CC acid from the medium. The vector of the invention is useful for

CC replacement of a wild-type pyruvate carboxylase gene, with a feedback

CC resistant pyruvate carboxylase gene, in a *Corynebacterium glutamicum*, by

CC replacing a genomic copy of the wild-type pyruvate carboxylase gene with

CC a selectable marker gene through homologous recombination to form a first

CC recombinant strain, and replacing the selectable marker gene in the

CC first recombinant strain, with feedback resistant pyruvate carboxylase

CC gene through homologous recombination to form a second recombinant

CC strain, where the homologous recombination in the above steps, occurs

CC between the host cell and the vector. The feedback-resistant pyruvate

CC carboxylase enzyme is resistant to feedback inhibition from aspartic

CC acid. The present amino acid sequence represents the mutant.

CC feedback-resistant pyruvate carboxylase enzyme of the invention.

CC Note: The present sequence is not shown in the specification but is

CC derived from the wild-type feedback-resistant pyruvate carboxylase

CC enzyme (AA098053) given in figure 2 of the specification.

CC

XX

SQ Sequence 1140 AA;

Query Match 44.3%; Score 151.5; DB 23; Length 1140;

Best Local Similarity 43.5%; Pred. No. 1.8e-09;

Matches 30; Conservative 15; Mismatches 23; Indels 1; Gaps 1;

OY 1 EGEIPLACTVSKILVKEGDTVAGQTVLLEAMKMETETINAPTDGKVEKLVKRDVAV 60

DB 1072 KGVNAPFAGCVT-VTVABDEVKAGDAVAITAMKMETITTSVSGKIDRVVPATKY 1130

OY 61 QGGGGLIKI 69

DB 1131 EGGDLIVV 1139

RESULT 22

ABP28014

XX ABP28014 standard; Protein; 166 AA.

XX ABP28014;

DT 02-JUL-2002 (first entry)

XX Streptococcus polypeptide SEQ ID NO 5204.

XX Streptococcus; GAS; group B streptococcus; Streptococcus agalactiae;

XX group A streptococcus; Streptococcus pyogenes; antibacterial;

XX anti-inflammatory; Infection; vaccine; meningitis; gene therapy.

XX Streptococcus pyogenes.

XX WC200234771-A2.

XX 02-MAY-2002.

XX 29-OCT-2001; 2001WO-GB04789.

XX 27-OCT-2000; 2000GB-0026333.

XX 24-NOV-2000; 2000GB-0028727.

XX 07-MAR-2001; 2001GB-0005640.

XX (CHIR-) CHIRON SPA.

XX (GENO-) INST GENOMIC RES.

XX Telford J, Masignani V, Margarit Ros YI, Grandi G, Fraser C;

XX Tettelin H;

XX WPI: 2002-352536/38.

XX N-PSDB; ABN68645.

PT New Streptococcus protein for the treatment or prevention of infection

PT or disease caused by Streptococcus bacteria, such as meningitis, and

PT for detecting a compound that binds to the protein -

PS Claim 1; Page 3684; 4525pp; English.

XX The invention relates to a protein (ABP25413-ABP30895) from group B

CC Streptococcus/GAS (Streptococcus agalactiae) or group A streptococcus/GAS

CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in

CC the specification. The proteins have antibacterial and anti-inflammatory

CC activity. (I), nucleic acids encoding (I), ABN6044-ABN71526 and

CC antibodies that bind (I) are used in the manufacture of medicaments for

CC the treatment or prevention of infection or disease caused by

CC Streptococcus bacteria, particularly *S. agalactiae* and *S. pyogenes*.

CC Nucleic acids encoding (I) are used to detect Streptococcus in a

CC biological sample. (I) is used to determine whether a compound binds to

CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be

CC used as a vaccine or diagnostic composition. The disease caused by

CC Streptococcus that is prevented or treated may be meningitis. Nucleic

CC acid encoding (I) may be used to recombinantly produce (I) and may be

CC used in gene therapy. Antibodies to (I) are used for affinity

CC chromatography, immunoassays, and distinguishing/identifying

CC Streptococcus proteins.

CC

XX

SQ Sequence 166 AA;

Query Match 44.2%; Score 151; DB 23; Length 166;

Best Local Similarity 41.6%; Pred. No. 1.7e-10;

Matches 32; Conservative 15; Mismatches 22; Indels 8; Gaps 2;

OY 1 EGEI-PAPLAGTV-----SKLVKEGDTVAGQTVLLEAMKMETETINAPTDGKVEV 52

DB 89 EGDIVSPLVGAVYLAASDPKPEFVAGDTVKKGDTLVIEAMKVNNEVPACDGVTEI 148

OY 53 LVKERDVOGGGLIKI 69

DB 149 LVSNEDVIERGGGLVRI 165

RESULT 23

AAU03126

XX AAU03126 standard; Protein; 134 AA.

XX AAU03126;

DT 23-OCT-2001 (first entry)

XX Streptococcus pyogenes virulence protein #8.

XX Virulence; attenuated microorganism; Streptococcal infection;

XX Gram-positive bacteria; antimicrobial; Impetigo; pneumonia.

XX Streptococcus pyogenes.

XX WO200148208-A2.

XX 05-JUL-2001.

XX 22-DEC-2000; 2000WO-GB04997.

XX 23-DEC-1999; 99GB-0030462.

XX 23-DEC-1999; 99GB-0030463.

XX 23-DEC-1999; 99GB-0030464.

XX 23-DEC-1999; 99GB-0030466.

XX 23-DEC-1999; 99GB-0030467.

XX 23-DEC-1999; 99GB-0030469.

XX 23-DEC-1999; 99GB-0030471.

XX 23-DEC-1999; 99GB-0030472.

XX 23-DEC-1999; 99GB-0030473.

XX 23-DEC-1999; 99GB-0030474.

XX 17-FEB-2000; 2000GB-0003725.

XX 17-FEB-2000; 2000GB-0003726.

XX 17-FEB-2000; 2000GB-0003727.

PR 17-FEB-2000: 2000GB-0003728.
PR 17-FEB-2000: 2000GB-0003729.
PR 17-FEB-2000: 2000GB-0003730.
PR 17-FEB-2000: 2000GB-0003731.
PR 17-FEB-2000: 2000GB-0003732.
PR 17-FEB-2000: 2000GB-0003733.
PR 02-MAY-2000: 2000GB-0010585.
PR 02-MAY-2000: 2000GB-0010587.
XX (MICR-) MICROSCIENCE LTD.
XX
PI Claire EE, Zhou L, Shea JE, Feldman RG, Holden DW;
XX WPI; 2001-418285/44.
DR N-PSDB; AAS06358.
XX
PT Novel peptide obtained from Streptococcus pyogenes useful for treating
PT or preventing a condition associated with infection by Streptococcal or
XX Gram-positive bacteria, preferably pneumonia
XX
PS Claim 4; Page 40-41; 91pp; English.
XX
CC AAU03119-AAU03149 represent novel Streptococcus pyogenes virulence
CC proteins #1-31. The S. pyogenes virulence genes can be used to
CC produce attenuated microorganisms comprising a mutation that disrupt
CC the expression of the virulence protein. The virulence genes, proteins
CC or an attenuated microorganism are useful for therapeutic or diagnostic
CC purposes. DNA encoding the virulence proteins, the proteins themselves,
CC an attenuated microorganism or a vaccine comprising the virulence
CC protein are useful for the manufacture of a medicament for use in
CC the treatment or prevention of a condition associated with infection
CC by Streptococcal or Gram-positive bacteria, for veterinary treatment,
CC and in a screening assay for the identification of an antimicrobial
CC drug. Disorders which can be treated using S. pyogenes virulence
CC polynucleotide and polypeptide sequences include non-invasive infections
CC and invasive infections e.g. impetigo, pharyngitis, necrotising
CC fasciitis, bacteraemia, streptococcal toxic shock syndrome (STSS),
CC pneumonia and rheumatic fever. The virulence proteins are also useful
CC in the preparation of antibodies.
XX
SQ Sequence 134 AA;
XX
Query Match 43.0%; Score 147; DB 22; Length 134;
Best Local Similarity 47.8%; Pred. No. 3.9e-10;
Matches 32; Conservative 9; Mismatches 26; Indels 0; Gaps 0;
XX
QY 4 IPAPLAGTAVSKILVKEGDTVAKGTVLLEAMKMETEINAPTDGKVERKLVKERDAYOGG 63
DB 66 IPSPMPTGIKLVYAVGDQYTENQPLILTEAMKMETEINAPTDGKVERKLVKERDAYOGG 125
XX
QY 64 QGLIKIG 70
DB 126 DGLITIG 132
XX
RESULT 24
AAU36768
ID AAU36768 standard; Protein; 1073 AA.
XX
AC AAU36768;
XX
DT 14-FEB-2002 (first entry)
XX
DE Staphylococcus aureus cellular proliferation protein #938.
XX
KM Antisense: prokaryotic cellular proliferation protein;
XX antibiotic; antibacterial; drug design.
XX
OS Staphylococcus aureus.
XX
PN WO200170955-A2.
XX
PD 27-SEP-2001.

XX
XX 21-MAR-2001: 2001MO-US09180.
XX
XX 21-MAR-2000: 2000US-191078P.
PR 23-MAY-2000: 2000US-206848P.
PR 26-MAY-2000: 2000US-207727P.
PR 23-OCT-2000: 2000US-242578P.
PR 27-NOV-2000: 2000US-253625P.
PR 22-DEC-2000: 2000US-257931P.
PR 16-FEB-2001: 2001US-269308P.
XX
XX (ELIR-) ELITRA PHARM INC.
XX
XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;
XX
DR WPI; 2001-611495/70.
DR N-PSDB; AAS54627.
XX
PT New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids -
XX
XX Example 3; Seq ID No 12361; 511pp; English.
XX
XX The invention relates to antisense inhibitors of genes essential to
XX prokaryotic cellular proliferation, their use in identifying the
XX genes, their use in the discovery of novel antibiotics, the essential
XX genes themselves and the encoded proteins. The prokaryotes used are
XX Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
XX pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
XX invention is also useful for the identification of potential new targets
XX for antibiotic development. The antisense nucleic acids can also be used
XX to identify proteins used in proliferation, to express these proteins,
XX and to obtain antibodies capable of binding to the expressed proteins.
XX The proteins can be used to screen compounds in rational drug discovery
XX programmes. The antisense nucleic acid sequence is also useful to screen
XX for homologous nucleic acids which are required for cell proliferation in
XX a wide variety of organisms. The present sequence represents an
XX essential prokaryotic cellular proliferation protein.
XX CC Note: The sequence data for this patent did not form part
XX of the printed specification, but was obtained in electronic
XX format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1073 AA;
XX
Query Match 42.7%; Score 146; DB 22; Length 1073;
Best Local Similarity 47.0%; Pred. No. 7.8e-09;
Matches 31; Conservative 13; Mismatches 22; Indels 0; Gaps 0;
XX
QY 4 IPAPLAGTAVSKILVKEGDTVAKGTVLLEAMKMETEINAPTDGKVERKLVKERDAYOGG 63
DB 1003 IGAQMPGSVTEVKVSGEYVAAQNPPLITTEAMKMETEINAPTDGKVERKLVKERDAYOGG 1062
XX
QY 64 QGLIKI 69
DB 1063 DLIRI 1068
XX
RESULT 25
AAU33972
ID AAU33972 standard; Protein; 1147 AA.
XX
AC AAU33972;
XX
DT 14-FEB-2002 (first entry)
XX
DE Staphylococcus aureus cellular proliferation protein #248.
XX
KM Antisense: prokaryotic cellular proliferation protein;
XX antibiotic; antibacterial; drug design.
XX
OS Staphylococcus aureus.

XX WO200170955-A2.
 PN
 XX 27-SEP-2001.
 PD
 XX
 PF 21-MAR-2001; 2001WO-US09180.
 XX
 XX 21-MAR-2000; 2000US-191078P.
 PR 23-MAY-2000; 2000US-206848P.
 PR 26-MAY-2000; 2000US-207727P.
 PR 23-OCT-2000; 2000US-242578P.
 PR 27-NOV-2000; 2000US-253625P.
 PR 22-DEC-2000; 2000US-257931P.
 PR 16-FEB-2001; 2001US-269308P.
 XX
 XX (ELIT-) ELITRA PHARM INC.
 PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
 PI Yamamoto RT, Xu HH;
 XX WPI: 2001-611495/70.
 DR N-PSDB; AAS51831.
 XX
 XX New polynucleotides for the identification and development of
 PT antibiotics, comprise sequences of antisense nucleic acids -
 XX
 PS Example 3; Seq ID No 5468; 511pp; English.
 XX
 CC The invention relates to antisense inhibitors of genes essential to
 CC prokaryotic cellular proliferation, their use in identifying the
 CC genes, their use in the discovery of novel antibiotics, the essential
 CC genes themselves and the encoded proteins. The prokaryotes used are
 CC *Escherichia coli*, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella*
 CC *pneumoniae*, *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The
 CC invention is also useful for the identification of potential new targets
 CC for antibiotic development. The antisense nucleic acids can also be used
 CC to identify proteins used in proliferation, to express these proteins,
 CC and to obtain antibodies capable of binding to the expressed proteins.
 CC The proteins can be used to screen compounds in rational drug discovery
 CC programmes. The antisense nucleic acid sequence is also useful to screen
 CC for homologous nucleic acids which are required for cell proliferation in
 CC a wide variety of organisms. The present sequence represents an
 CC essential prokaryotic cellular proliferation protein.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pcl_sequences.
 CC
 XX
 SQ Sequence 1147 AA;
 Query Match 42.7%; Score 146; DB 22; Length 1147;
 Best Local Similarity 47.0%; Pred. No. 8.5e-09;
 Matches 31; Conservative 13; Mismatches 22; Indels 0; Gaps 0;
 OY 4 IPAPLAGTSTKLVKRGDTVLEAKMKEINAPDGVKLVKENDAVOGG 63
 DB 1079 IGAKPMGSTVEKVSGETVYKANQPLITEAKMETTTIAPDGVKIVVNNGTIATG 1138
 OY 64 GGLIKI 69
 DB 1139 DLLIEI 1144
 RESULT 26
 ID AAY14912
 XX AAY14912 standard; protein; 108 AA.
 AC AAY14912;
 XX 25-OCT-1999 (first entry)
 DT M. vaccae antigen GV-29 3' sequence.
 DE
 XX

KW Mycobacterium vaccae protein; antigen; T cell activation; cytokine;
 KW dendritic cell maturation; infectious disease; immune disorder; cancer;
 KW respiratory system; mycobacterial infection; allergy; tuberculosis;
 KW leprosy; sarcoidosis; lung cancer; asthma; skin disorder; psoriasis;
 KW dermatitis; eczema; alopecia areata; skin cancer; basal carcinoma;
 KW squamous cell carcinoma; melanoma.
 XX
 OS Mycobacterium vaccae.
 XX
 XX WO9932634-A2.
 PN
 XX
 XX 01-JUL-1999.
 PD
 XX
 XX 23-DEC-1998; 98WO-NZ00189.
 PF
 XX
 XX 04-DEC-1998; 98US-0205426.
 PR 23-DEC-1997; 97US-0996624.
 PR 23-DEC-1997; 97US-0997080.
 PR 23-DEC-1997; 97US-0997362.
 PR 11-JUN-1998; 98US-0095855.
 PR 17-SEP-1998; 98US-0156181.
 XX
 PA (GENE-) GENESIS RES & DEV CORP LTD.
 XX
 PI Prestidge RL, Skinner MA, Tan P, Visser ES, Watson J;
 PI WPI: 1999-430163/36.
 DR N-PSDB; AA211374.
 XX
 XX Enhancing immune response to an antigen
 PS Claim 1; Page 217; 243pp; English.
 XX
 CC The invention provides heat-killed Mycobacterium vaccae, or recombinant
 CC M. vaccae proteins. The M. vaccae proteins may be employed to activate
 CC T cells and natural killer cells, to stimulate the production of
 CC cytokines, to enhance the expression of co-stimulatory molecules on
 CC dendritic cells and monocytes, and to enhance dendritic cell maturation
 CC and function. The proteins can be expressed by standard recombinant
 CC methodology. Pharmaceutical compositions comprising the proteins or
 CC nucleic acid sequences encoding the proteins can be used for the
 CC treatment, prevention, and detection of disorders including infectious
 CC diseases, immune disorders and cancer. In particular, the compounds and
 CC methods are used for treatment of diseases of the respiratory system,
 CC such as mycobacterial infections, asthma, allergies, tuberculosis,
 CC leprosy, sarcoidosis and lung cancers, and disorders of the skin such as
 CC psoriasis, atopic dermatitis, eczema, allergic contact dermatitis,
 CC alopecia areata, and skin cancers such as basal carcinoma, squamous cell
 CC carcinoma and melanoma.
 CC
 XX
 SQ Sequence 108 AA;
 Query Match 42.1%; Score 144; DB 20; Length 108;
 Best Local Similarity 47.8%; Pred. No. 6.9e-10;
 Matches 32; Conservative 9; Mismatches 26; Indels 0; Gaps 0;
 OY 4 IPAPLAGTSTKLVKRGDTVLEAKMKEINAPDGVKLVKENDAVOGG 63
 DB 37 VDAPFASSVYKVDVAVGDRVAGQPLLEAKMKEIVLAPADGVVTOIVSAGHLVDPG 96
 OY 64 GGLIKIG 70
 DB 97 TPLVVG 103
 RESULT 27
 ID ABB73518
 XX ABB73518 standard; protein; 108 AA.
 AC ABB73518;
 XX 08-APR-2002 (first entry)
 DT
 XX

DE M vaccae GV-29 partial protein SEQ ID NO: 166.
XX
KW Skin disorder; psoriasis; atopic dermatitis; allergic contact dermatitis;
KW alopecia areata; skin cancer; Mycobacterium vaccae; melanoma; cytostatic;
KW antiproliferative; dermatological; antiinflammatory; antiallergic;
KW Th2 immune response; immunomodulatory.
XX
OS Mycobacterium vaccae.
XX
PN US6328978-B1.
XX
PD 11-DEC-2001.
XX
PF 02-JUN-1999; 99US-0324542.
XX
PR 23-DEC-1997; 97US-0997080.
XX
PA (GENE-) GENESIS RES & DEV CORP LTD.
XX
PI Watson JD, Tan PJ, Prestidge R;
XX
DR WPI: 2002-138361/18.
XX
DR N-PSDB; ABL36280.
XX
PT Inhibiting skin inflammation associated with skin disorder e.g.
PT psoriasis, by administering composition comprising delipidated and
PT delipidated Mycobacterium vaccae cells or Mycobacterium vaccae
PT culture filtrate
XX
PS Example 4; Column 177-178; 116pp; English.
XX
CC The present invention relates to a method of inhibiting skin inflammation
CC associated with a skin disorder selected from psoriasis, atopic
CC dermatitis and allergic contact dermatitis, which involves administering
CC a composition containing delipidated and delipidated Mycobacterium
CC vaccae cells or M. vaccae culture filtrate. The skin disorder to be
CC treated may also include alopecia areata, and skin cancers such as basal
CC cell carcinoma, squamous cell carcinoma and melanoma. The composition
CC acts by inhibiting the Th2 immune response. The present sequence is a
CC protein described in the exemplification of the invention.
XX
SQ Sequence 108 AA;
SO
Query Match 42.1%; Score 144; DB 23; Length 108;
Best Local Similarity 47.8%; Pred. No. 6.9e-10;
Matches 32; Conservative 9; Mismatches 26; Indels 0; Gaps 0;
QY 4 IPAPLAGTVSKILVKEGDTYKAGOTVLEAMKMETEINAPTDGVEKVKERDAVGG 63
DB 37 VDAPFASVWVYDAVGVVAGPPLALEAMKMETVLRAPADGVVTOILVSAGHLVDPG 96
QY 64 OGLIKIG 70
DB 97 TPLVYVG 103
RESULT 28
AAU37770
ID AAU37770 standard; Protein: 161 AA.
XX
AC AAU37770;
XX
DT 14-FEB-2002 (first entry)
XX
DE Streptococcus pneumoniae cellular proliferation protein #199.
XX
KW Antisense; prokaryotic cellular proliferation protein;
KW antibiotic; antibacterial; drug design.
XX
OS Streptococcus pneumoniae.
XX
PN WO200170955-A2.
XX

PD 27-SEP-2001.
XX
PF 21-MAR-2001; 2001WO-US09180.
XX
PR 21-MAR-2000; 2000US-191078P.
PR 23-MAY-2000; 2000US-206848P.
PR 26-MAY-2000; 2000US-207127P.
PR 23-OCT-2000; 2000US-242578P.
PR 27-NOV-2000; 2000US-253625P.
PR 22-DEC-2000; 2000US-257931P.
PR 16-FEB-2001; 2001US-269308P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Haselbeck R, Ohlsen KL, Zykkind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;
XX
DR WPI: 2001-611495/70.
XX
DR N-PSDB; AAS55629.
XX
PT New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids -
XX
XX
PS Example 3; Seq ID No 13363; 51pp; English.
XX
CC The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the
CC genes, their use in the discovery of novel antibiotics, the essential
CC genes themselves and the encoded proteins. The prokaryotes used are
CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC invention is also useful for the identification of potential new targets
CC for antibiotic development. The antisense nucleic acids can also be used
CC to identify proteins used in proliferation, to express these proteins,
CC and to obtain antibodies capable of binding to the expressed proteins.
CC The proteins can be used to screen compounds in rational drug discovery
CC programmes. The antisense nucleic acid sequence is also useful to screen
CC for homologous nucleic acids which are required for cell proliferation in
CC a wide variety of organisms. The present sequence represents an
CC essential prokaryotic cellular proliferation protein.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 161 AA;
SO
Query Match 42.1%; Score 144; DB 22; Length 161;
Best Local Similarity 41.6%; Pred. No. 1.2e-09;
Matches 32; Conservative 15; Mismatches 22; Indels 8; Gaps 2;
QY 1 EGE-IPAPLAGTV-----SKILVKEGDTYKAGOTVLEAMKMETEINAPTDGVEK 52
DB 84 EGNLVESPLVGVVYLAAGPDKPAFYVGSYKGGTLYIIRAMKYNIRPAKGVYTEI 143
QY 53 LVKERDAVGGGGLIKI 69
DB 144 LVSNEEMVEFGGLVRI 160
RESULT 29
AAU38022
ID AAU38022 standard; Protein: 161 AA.
XX
AC AAU38022;
XX
DT 14-FEB-2002 (first entry)
XX
DE Streptococcus pneumoniae cellular proliferation protein #451.
XX
KW Antisense; prokaryotic cellular proliferation protein;
KW antibiotic; antibacterial; drug design.
XX

OS Streptococcus pneumoniae.
 XX
 PN W0200170955-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 21-MAR-2001; 2001WO-US09180.
 XX
 PR 21-MAR-2000; 2000US-191078P.
 PR 23-MAY-2000; 2000US-206848P.
 PR 26-MAY-2000; 2000US-207727P.
 PR 23-OCT-2000; 2000US-242578P.
 PR 27-NOV-2000; 2000US-253625P.
 PR 22-DEC-2000; 2000US-257931P.
 PR 16-FEB-2001; 2001US-269308P.
 XX
 PA (ELITR-) ELITRA PHARM INC.
 XX
 PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ,
 PI Yamamoto RT, Xu HH;
 XX
 DR WPI; 2001-611495/70.
 DR N-PSDB; AAS55881.
 XX
 PT New polynucleotides for the identification and development of
 PT antibiotics, comprise sequences of antisense nucleic acids -
 XX
 PS Example 3; Seq ID No 13615; 511pp; English.
 XX
 CC The invention relates to antisense inhibitors of genes essential to
 CC prokaryotic cellular proliferation, their use in identifying the
 CC genes, their use in the discovery of novel antibiotics, the essential
 CC genes themselves and the encoded proteins. The prokaryotes used are
 CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
 CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
 CC invention is also useful for the identification of potential new targets
 CC for antibiotic development. The antisense nucleic acids can also be used
 CC to identify proteins used in proliferation, to express these proteins,
 CC and to obtain antibodies capable of binding to the expressed proteins.
 CC The proteins can be used to screen compounds in rational drug discovery
 CC programmes. The antisense nucleic acid sequence is also useful to screen
 CC for homologous nucleic acids which are required for cell proliferation in
 CC a wide variety of organisms. The present sequence represents an
 CC essential prokaryotic cellular proliferation protein.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 161 AA:
 Query Match 42.1%; Score 144; DB 22; Length 161;
 Best Local Similarity 41.6%; Pred. No. 1.2e-09;
 Matches 32; Conservative 15; Mismatches 22; Indels 8; Gaps 2;
 QY 1 EGE-IPAPLAGIV-----SKLVKESDITVRAGQTVLVLAMKMETETINAPDGVKEY 52
 DB 84 EGNLVESPLVGVVYLAAGDPKPAFVTVGDSVKKGGQTVLITAMKVNELIPAPDGVVTEI 143
 QY 53 LYKERDVAOGGGLIKI 69
 DB 144 LVSNEEMVEFGKGLVRI 160
 RESULT 30
 ID AAM01086
 ID AAM01086 standard; Protein; 161 AA.
 AC AAM01086;
 AC
 DT 02-OCT-2001 (first entry)
 XX
 DE CFE 09 protein sequence.

XX Antibacterial; vaccine; gene therapy; bacterial cell wall viability;
 KM CFE; Conserved Essential Gene; bacterial infection;
 KM antisense therapy; antibiotic resistance.
 XX
 OS Streptococcus pneumoniae.
 XX
 PN W0200149721-A2.
 XX
 PD 12-JUL-2001.
 XX
 PR 29-DEC-2000; 2000WO-US35604.
 PR 30-DEC-1999; 99US-0174089.
 XX
 PA (BRIM) BRISTOL-MYERS SQUIBB CO.
 XX
 PI Dougherty TJ, Pucci MJ, Dougherty BA, Davison DB, Bruccoleri RE,
 PI Thanassi JA;
 XX
 DR WPI; 2001-496721/54.
 DR N-PSDB; AAM90785.
 XX
 PT Nucleic acids encoding conserved essential genes involved in bacterial
 PT replication which are potential targets for the treatment of antibiotic
 PT resistant bacterial infections -
 XX
 PS Claim 27; Page 333; 380pp; English.
 XX
 CC The present invention relates to nucleic acids (AAM90701-AAM90918)
 CC encoding polypeptides (AAM01002-AAM01114), which are essential for the
 CC viability of a bacterial cell wall. The acronym CFE stands for "Conserved
 CC Expression", where CFE stands for "Conserved Essential Gene". The nucleic
 CC acids are useful for detecting the presence of proteins essential for the
 CC viability of a bacterial cell wall in samples such as cells, tissues,
 CC biological fluids, blood, serum, nose, ear or throat swabs with ligands,
 CC and for detecting corresponding target nucleic acid molecules with
 CC complementary sequences. The nucleic acids are also useful for
 CC determining whether a genomic nucleotide sequence of interest is
 CC essential for viability of a bacterial cell or whether it resides within
 CC an operon, by integrating an exogenous nucleotide sequence comprising a
 CC portion of an open reading frame of the genomic sequence of interest
 CC (comprising 200-500 base pairs) into the genomic sequence of interest
 CC which confers a selectable phenotype to the cell, and determining cell
 CC viability with a selection agent such as chloramphenicol. The nucleic
 CC acids and proteins are also useful as vaccines and for treating bacterial
 CC infections with gene therapy and antisense therapy. The nucleic acids
 CC also enable identification of targets suitable for the treatment of
 CC antibiotic resistant bacterial infections.
 XX
 SQ Sequence 161 AA:
 Query Match 42.1%; Score 144; DB 22; Length 161;
 Best Local Similarity 41.6%; Pred. No. 1.2e-09;
 Matches 32; Conservative 15; Mismatches 22; Indels 8; Gaps 2;
 QY 1 EGE-IPAPLAGIV-----SKLVKESDITVRAGQTVLVLAMKMETETINAPDGVKEY 52
 DB 84 EGNLVESPLVGVVYLAAGDPKPAFVTVGDSVKKGGQTVLITAMKVNELIPAPDGVVTEI 143
 QY 53 LYKERDVAOGGGLIKI 69
 DB 144 LVSNEEMVEFGKGLVRI 160

Search completed: May 1, 2003, 07:53:55
 Job time : 23.6042 secs

ALIGNMENTS

RESULT 1
US-09-364-083-2

Sequence 2, Application US/09364083
Patent No. 6210676
GENERAL INFORMATION:
APPLICANT: Callister, Steven M
APPLICANT: Lovrich, Steven D
APPLICANT: Schell, Ronald F

APPLICANT: Jobe, Dean A
TITLE OF INVENTION: Compositions and Method Using the Borrelliacidal
TITLE OF INVENTION: Epitope(s) of Borrelia burgdorferi Outer Surface
TITLE OF INVENTION: Protein C (OspC) for the diagnosis and prevention of
FILE REFERENCE: B. burgdorferi OspC
CURRENT APPLICATION NUMBER: US/09/364,083
CURRENT FILING DATE: 1999-07-30
EARLIER APPLICATION NUMBER: 60/094,955
EARLIER FILING DATE: 1998-07-31
NUMBER OF SEQ ID NOS: 4
SOFTWARE: Patentin Ver. 2.0

SEQ ID NO 2
LENGTH: 194
TYPE: PRT
ORGANISM: Borrelia burgdorferi
US-09-364-083-2

Query Match 100.0%; Score 615; DB 4; Length 194;
Best Local Similarity 100.0%; Pred. No. 1.7e-57;
Matches 122; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKLKVTVNGTAYDVVDVDSKSHENPMGTLIFGGGTGAPAPAAAGAGAGAGGEGTIPAPL 60
DB 1 MKLKVTVNGTAYDVVDVDSKSHENPMGTLIFGGGTGAPAPAAAGAGAGAGGEGTIPAPL 60
QY 61 AGTVSKILVKEGDTVKAGQTVLVLEAMKMETEINAPTDGKVEKVLVKERDAVVOGGGLIK 120
DB 61 AGTVSKILVKEGDTVKAGQTVLVLEAMKMETEINAPTDGKVEKVLVKERDAVVOGGGLIK 120
QY 121 IG 122
DB 121 IG 122

RESULT 2
US-09-651-419-2

Sequence 2, Application US/09651419
Patent No. 6464985
GENERAL INFORMATION:
APPLICANT: Callister, Steven M
APPLICANT: Lovrich, Steven D
APPLICANT: Schell, Ronald F

APPLICANT: Jobe, Dean A
TITLE OF INVENTION: Compositions and Method Using the Borrelliacidal
TITLE OF INVENTION: Epitope(s) of Borrelia burgdorferi Outer Surface
TITLE OF INVENTION: Protein C (OspC) for the diagnosis and prevention of
FILE REFERENCE: B. burgdorferi OspC
CURRENT APPLICATION NUMBER: US/09/651,419
CURRENT FILING DATE: 2000-08-30
PRIOR APPLICATION NUMBER: 60/094,955
PRIOR FILING DATE: 1998-07-31
NUMBER OF SEQ ID NOS: 4
SOFTWARE: Patentin Ver. 2.0

SEQ ID NO 2
LENGTH: 194
TYPE: PRT
ORGANISM: Borrelia burgdorferi
US-09-651-419-2

Query Match 100.0%; Score 615; DB 4; Length 194;
Best Local Similarity 100.0%; Pred. No. 1.7e-57;
Matches 122; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKLKVTVNGTAYDVVDVDSKSHENPMGTLIFGGGTGAPAPAAAGAGAGAGGEGTIPAPL 60
DB 1 MKLKVTVNGTAYDVVDVDSKSHENPMGTLIFGGGTGAPAPAAAGAGAGAGGEGTIPAPL 60
QY 61 AGTVSKILVKEGDTVKAGQTVLVLEAMKMETEINAPTDGKVEKVLVKERDAVVOGGGLIK 120
DB 61 AGTVSKILVKEGDTVKAGQTVLVLEAMKMETEINAPTDGKVEKVLVKERDAVVOGGGLIK 120
QY 121 IG 122
DB 121 IG 122

RESULT 3

US-07-956-700B-107
Sequence 107, Application US/07956700B

Patent No. 5530922
GENERAL INFORMATION:

APPLICANT: Robert Haselkorn and Piotr Gornicki
TITLE OF INVENTION: Cyanobacterial and Plant Acetyl-CoA
TITLE OF INVENTION: Carboxylase
NUMBER OF SEQUENCES: 116
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: 321 No. 5530922th Clark Street
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60610

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy Disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII-DOS
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/956,700B
FILING DATE: 19921002

CLASSIFICATION: 800

ATTORNEY/AGENT INFORMATION:
NAME: Thomas E. No. 5530922thrup
REGISTRATION NUMBER: 33,268
REFERENCE/DOCKET NUMBER: ARCD:058
TELECOMMUNICATION INFORMATION:
TELEPHONE: 1-312-744-0090
TELEFAX: 1-312-755-4489

INFORMATION FOR SEQ ID NO: 107:

SEQUENCE CHARACTERISTICS:
LENGTH: 123 amino acids
TYPE: amino acid
STRANDEDNESS: Single
TOPOLOGY: linear
MOLECULE TYPE: Peptide
US-07-956-700B-107

Query Match 98.3%; Score 604.5; DB 1; Length 123;
Best Local Similarity 99.2%; Pred. No. 1.2e-56;
Matches 122; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 MKLKVTVNGTAYDVVDVDSKSHENPMGTLIFGGGTGAPAPAAAGAGAGAGGEGTIPAPL 59
DB 1 MKLKVTVNGTAYDVVDVDSKSHENPMGTLIFGGGTGAPAPAAAGAGAGAGGEGTIPAPL 60
QY 60 LAGTVSKILVKEGDTVKAGQTVLVLEAMKMETEINAPTDGKVEKVLVKERDAVVOGGGLI 119
DB 61 LAGTVSKILVKEGDTVKAGQTVLVLEAMKMETEINAPTDGKVEKVLVKERDAVVOGGGLI 120
QY 120 KIG 122
DB 120 KIG 122

Db 121 KIG 123

RESULT 4

US-08-476-537-107

Sequence 107 Application US/08476537

Patent No. 5756290

GENERAL INFORMATION:

APPLICANT: Robert Haselkorn and Piotr Gornicki

TITLE OF INVENTION: Cyanobacterial and Plant Acetyl-CoA

TITLE OF INVENTION: Carboxylase

NUMBER OF SEQUENCES: 116

CORRESPONDENCE ADDRESS:

ADDRESSEE: Arnold, White & Durkee

STREET: 321 No. 5756290th Clark Street

CITY: Chicago

STATE: Illinois

COUNTRY: USA

ZIP: 60610

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy Disk

COMPUTER: IBM PC Compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: ASCII-DOS

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/476,537

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/956,700

FILING DATE: 10/21/92

ATTORNEY/AGENT INFORMATION:

NAME: Thomas E. No. 5756290thrup

REGISTRATION NUMBER: 33,268

REFERENCE/DOCKET NUMBER: ARCD:058

TELECOMMUNICATION INFORMATION:

TELEPHONE: 1-312-744-0090

TELEFAX: 1-312-755-4489

INFORMATION FOR SEQ ID NO: 107:

SEQUENCE CHARACTERISTICS:

LENGTH: 123 amino acids

TYPE: Amino acid

STRANDEDNESS: Single

TOPOLOGY: Linear

MOLECULE TYPE: Peptide

US-08-476-537-107

Query Match

Best Local Similarity 98.3%; Score 604.5; DB 1; Length 123;

Matches 122; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Db 1 MKLKYVNGTAVDVVDVDRKSHENPMGTLFFGGTGAGAPAP-AAGGAGGKAGBEIRAP 59

Db 1 MKLKYVNGTAVDVVDVDRKSHENPMGTLFFGGTGAGAPAPRAAGAGAGKAGBEIRAP 60

Db 60 LAGTVSKILVKGEDTVKAGQVTLVLEAMKMETEINAPTDGKYEKYLVERDAVGGGGLI 119

Db 61 LAGTVSKILVKGEDTVKAGQVTLVLEAMKMETEINAPTDGKYEKYLVERDAVGGGGLI 120

OY 120 KIG 122

Db 121 KIG 123

RESULT 5

US-08-485-607-107

Sequence 107 Application US/08485607

Patent No. 5792627

GENERAL INFORMATION:

APPLICANT: Robert Haselkorn and Piotr Gornicki

TITLE OF INVENTION: Cyanobacterial and Plant Acetyl-CoA

TITLE OF INVENTION: Carboxylase

NUMBER OF SEQUENCES: 116

CORRESPONDENCE ADDRESS:

ADDRESSEE: Arnold, White & Durkee

STREET: 321 No. 5756290th Clark Street

CITY: Chicago

STATE: Illinois

COUNTRY: USA

ZIP: 60610

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy Disk

COMPUTER: IBM PC Compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

CORRESPONDENCE ADDRESS:

ADDRESSEE: Arnold, White & Durkee

STREET: 321 No. 5792627th Clark Street

CITY: Chicago

STATE: Illinois

COUNTRY: USA

ZIP: 60610

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy Disk

COMPUTER: IBM PC Compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: ASCII-DOS

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/485,607

FILING DATE: 07-JUN-1995

CLASSIFICATION: 800

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/956,700

FILING DATE: 10/21/92

ATTORNEY/AGENT INFORMATION:

NAME: Thomas E. No. 5792627thrup

REGISTRATION NUMBER: 33,268

REFERENCE/DOCKET NUMBER: ARCD:058

TELECOMMUNICATION INFORMATION:

TELEPHONE: 1-312-744-0090

TELEFAX: 1-312-755-4489

INFORMATION FOR SEQ ID NO: 107:

SEQUENCE CHARACTERISTICS:

LENGTH: 123 amino acids

TYPE: Amino acid

STRANDEDNESS: Single

TOPOLOGY: Linear

MOLECULE TYPE: Peptide

US-08-485-607-107

Query Match

Best Local Similarity 98.3%; Score 604.5; DB 1; Length 123;

Matches 122; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Db 1 MKLKYVNGTAVDVVDVDRKSHENPMGTLFFGGTGAGAPAP-AAGGAGGKAGBEIRAP 59

Db 1 MKLKYVNGTAVDVVDVDRKSHENPMGTLFFGGTGAGAPAPRAAGAGAGKAGBEIRAP 60

Db 60 LAGTVSKILVKGEDTVKAGQVTLVLEAMKMETEINAPTDGKYEKYLVERDAVGGGGLI 119

Db 61 LAGTVSKILVKGEDTVKAGQVTLVLEAMKMETEINAPTDGKYEKYLVERDAVGGGGLI 120

OY 120 KIG 122

Db 121 KIG 123

RESULT 6

US-08-475-879-107

Sequence 107 Application US/08475879

Patent No. 5972644

Patent No. 5972644 5786170

GENERAL INFORMATION:

APPLICANT: Robert Haselkorn and Piotr Gornicki

TITLE OF INVENTION: Cyanobacterial and Plant Acetyl-CoA

TITLE OF INVENTION: Carboxylase

NUMBER OF SEQUENCES: 116

CORRESPONDENCE ADDRESS:

ADDRESSEE: Arnold, White & Durkee

STREET: 321 No. 5972644 5786170th Clark Street

CITY: Chicago

STATE: Illinois

COUNTRY: USA

ZIP: 60610

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy Disk

COMPUTER: IBM PC Compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: ASCII-DOS
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/4/5, 879
FILING DATE: 07-JUN-1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/956,700
FILING DATE: 10/21/92
ATTORNEY/AGENT INFORMATION:
NAME: Thomas E. No. 5972644 5786170chrup
REGISTRATION NUMBER: 33,268
REFERENCE/DOCKET NUMBER: ARCD:058
TELECOMMUNICATION INFORMATION:
TELEPHONE: 1-312-744-0090
TELEFAX: 1-312-755-4489
INFORMATION FOR SEQ ID NO: 107:
SEQUENCE CHARACTERISTICS:
LENGTH: 123 amino acids
TYPE: Amino acid
STRANDEDNESS: Single
TOPOLOGY: Linear
MOLECULE TYPE: Peptide
US-08-475-879-107

Query Match 98.3%; Score 604.5; DB 2; Length 123;
Best Local Similarity 99.2%; Pred. No. 1.2e-56;
Matches 122; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 MKLKVNGTAYDVVDVVDKSHENPMGTILFGGTTGAPAP-AAGGAGKAGEGEPAP 59
DB 1 MKLKVNGTAYDVVDVVDKSHENPMGTILFGGTTGAPAPRAAGAGKAGEGEPAP 60
QY 60 LACTVSKILVKEGDTYKAGQTVLVLEAMKMETEINAPTDGKVEKVLKERDAVQGGGLI 119
DB 61 LACTVSKILVKEGDTYKAGQTVLVLEAMKMETEINAPTDGKVEKVLKERDAVQGGGLI 120
QY 120 KIG 122
DB 121 KIG 123

RESULT 7
US-07-687-819-1
Sequence 1, Application US/07687819
Patent No. 6072039
GENERAL INFORMATION:
APPLICANT: Haase, Ferdinand C.
APPLICANT: Crease, Dean E.
TITLE OF INVENTION: Carrier Protein
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Rohm and Haas Company
STREET: Independence Mall West
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19105
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/687,819
FILING DATE: 19910520
CLASSIFICATION: 530
CLASSIFICATION: C07K13/00
CLASSIFICATION: C07K15/26
CLASSIFICATION: C12P21/00
CLASSIFICATION: 424/85.7
CLASSIFICATION: 530/351
CLASSIFICATION: 435/69.51
ATTORNEY/AGENT INFORMATION:

NAME: Dricks, Jordan J.
REGISTRATION NUMBER: 22,029
REFERENCE/DOCKET NUMBER: 89-020
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-592-2478
TELEFAX: 215-592-2682
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 123 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: *Protonibacterium shermanii*
PUBLICATION INFORMATION:
AUTHORS: Maloy, W L
AUTHORS: Bowlen, B U
AUTHORS: Zwolinski, G K
AUTHORS: Kumar, K G
AUTHORS: Wood, H G
TITLE: Amino Acid Sequence of the Biotinyl Subunit
Patent No. 6072039
JOURNAL: Journal of Biological Chemistry
VOLUME: 254
ISSUE: 22
PAGES: 11615-11622
DATE: No. 6072039 26-1979
RELEVANT RESIDUES IN SEQ ID NO: 1: FROM 1 TO 123
US-07-687-819-1

Query Match 98.3%; Score 604.5; DB 3; Length 123;
Best Local Similarity 99.2%; Pred. No. 1.2e-56;
Matches 122; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 MKLKVNGTAYDVVDVVDKSHENPMGTILFGGTTGAPAP-AAGGAGKAGEGEPAP 59
DB 1 MKLKVNGTAYDVVDVVDKSHENPMGTILFGGTTGAPAPRAAGAGKAGEGEPAP 60
QY 60 LACTVSKILVKEGDTYKAGQTVLVLEAMKMETEINAPTDGKVEKVLKERDAVQGGGLI 119
DB 61 LACTVSKILVKEGDTYKAGQTVLVLEAMKMETEINAPTDGKVEKVLKERDAVQGGGLI 120
QY 120 KIG 122
DB 121 KIG 123

RESULT 8
US-09-433-043B-107
Sequence 107, Application US/09433043B
Patent No. 6399342
GENERAL INFORMATION:
APPLICANT: HASELKORN, ROBERT
APPLICANT: GORNIK, PIOTR
TITLE OF INVENTION: CYANOBACTERIAL AND PLANT ACETYL-COA CARBOXYLASE
FILE REFERENCE: ARCD:33805
CURRENT APPLICATION NUMBER: US/09/433,043B
CURRENT FILING DATE: 1999-10-25
PRIOR APPLICATION NUMBER: 08/475,879
PRIOR FILING DATE: 1995-06-07
PRIOR APPLICATION NUMBER: 07/956,700
PRIOR FILING DATE: 1992-10-02
NUMBER OF SEQ ID NOS: 128
SOFTWARE: Patent Ver. 2.1
SEQ ID NO 107
LENGTH: 123
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-433-043B-107

FILING DATE: 19910520
CLASSIFICATION: 530
CLASSIFICATION: C07K13/00
CLASSIFICATION: C07K15/26
CLASSIFICATION: C12P21/00
CLASSIFICATION: 424/85.7
CLASSIFICATION: 530/351
CLASSIFICATION: 435/69.51
ATTORNEY/AGENT INFORMATION:
NAME: Drilks, Jordan J.
REGISTRATION NUMBER: 22,029
REFERENCE/DOCKET NUMBER: 89-020
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-592-2478
TELEFAX: 215-592-2682
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 43 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: Internal
US-07-687-819-2

Query Match 34.3%; Score 211; DB 3; Length 43;
Best Local Similarity 100.0%; Pred. No. 9.9e-16;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 57 PAPLAGTAVSKILVKEGDTVAKGOTVILVLEAMKMETEINAPTDG 99
DB 1 PAPLAGTAVSKILVKEGDTVAKGOTVILVLEAMKMETEINAPTDG 43

RESULT 13
US-09-433-043B-127
Sequence 127, Application US/09433043B
Patent No. 6399342
GENERAL INFORMATION:
APPLICANT: HASELKORN, ROBERT
APPLICANT: GORNICKI, PIOTR
TITLE OF INVENTION: CYANOBACTERIAL AND PLANT ACETYL-COA CARBOXYLASE
FILE REFERENCE: ARCD:338US
CURRENT APPLICATION NUMBER: US/09/433,043B
CURRENT FILING DATE: 1999-10-25
PRIOR APPLICATION NUMBER: 08/475,879
PRIOR FILING DATE: 1995-06-07
PRIOR APPLICATION NUMBER: 07/956,700
PRIOR FILING DATE: 1992-10-02
NUMBER OF SEQ ID NOS: 128
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 127
LENGTH: 170
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
OTHER INFORMATION: Peptide
US-09-433-043B-127

Query Match 32.9%; Score 202.5; DB 4; Length 170;
Best Local Similarity 52.9%; Pred. No. 4.4e-14;
Matches 45; Conservative 10; Mismatches 29; Indels 1; Gaps 1;

OY 38 AAPAPAGGAGKAGGEG-EIPAPLAGTAVSKILVKEGDTVAKGOTVILVLEAMKMETEINAP 96
DB 85 AAPAPASAPAAAAPAGAGPVTAPLAGTIVKVLASGOTVAAAGEVILLILEAMKMETEIRAA 144
OY 97 TDGKVEKVLVKERDAVVGSGGLIKI 121
DB 145 QAGTVRGIAVAKGADAVAVGDTLMTL 169

RESULT 14

US-07-956-700B-106
Sequence 106, Application US/07956700B
Patent No. 5539092
GENERAL INFORMATION:
APPLICANT: Robert Haselkorn and Piotr Gornicki
TITLE OF INVENTION: Cyanobacterial and Plant Acetyl-CoA
NUMBER OF SEQUENCES: 116
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: 321 No. 5539092th Clark Street
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60610
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy Disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII-DOS
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/956,700B
FILING DATE: 19921002
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Thomas E. No. 5539092thrup
REGISTRATION NUMBER: 33,268
REFERENCE/DOCKET NUMBER: ARCD:058
TELECOMMUNICATION INFORMATION:
TELEPHONE: 1-312-744-0090
TELEFAX: 1-312-755-4489
INFORMATION FOR SEQ ID NO: 106:
SEQUENCE CHARACTERISTICS:
LENGTH: 165 amino acids
TYPE: Amino acid
STRANDEDNESS: Single
TOPOLOGY: Linear
MOLECULE TYPE: Peptide
US-07-956-700B-106

Query Match 32.6%; Score 200.5; DB 1; Length 165;
Best Local Similarity 52.9%; Pred. No. 6.9e-14;
Matches 45; Conservative 10; Mismatches 29; Indels 1; Gaps 1;

OY 38 AAPAPAGGAGKAGGEG-EIPAPLAGTAVSKILVKEGDTVAKGOTVILVLEAMKMETEINAP 96
DB 80 AAPAPASAPAAAAPAGAGPVTAPLAGTIVKVLASGOTVAAAGEVILLILEAMKMETEIRAA 139
OY 97 TDGKVEKVLVKERDAVVGSGGLIKI 121
DB 140 QAGTVRGIAVAKGADAVAVGDTLMTL 164

RESULT 15
US-08-476-537-106
Sequence 106, Application US/08476537
Patent No. 5756290
GENERAL INFORMATION:
APPLICANT: Robert Haselkorn and Piotr Gornicki
TITLE OF INVENTION: Cyanobacterial and Plant Acetyl-CoA
NUMBER OF SEQUENCES: 116
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: 321 No. 5756290th Clark Street
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60610
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy Disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

* SOFTWARE: ASCII-DOS
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/476,537
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/956,700
FILING DATE: 10/21/92
ATTORNEY/AGENT INFORMATION:
NAME: Thomas E. No. 5756290thrup
REGISTRATION NUMBER: 33,268
REFERENCE/DOCKET NUMBER: ARCD:058
TELECOMMUNICATION INFORMATION:
TELEPHONE: 1-312-744-0090
TELEFAX: 1-312-755-4489
INFORMATION FOR SEQ ID NO: 106:
SEQUENCE CHARACTERISTICS:
LENGTH: 165 amino acids
TYPE: Amino acid
STRANDEDNESS: Single
TOPOLOGY: Linear
MOLECULE TYPE: Peptide
US-08-476-537-106

Query Match 32.6%; Score 200.5; DB 1; Length 165;
Best Local Similarity 52.9%; Pred. No. 6.9e-14;
Matches 45; Conservative 10; Mismatches 29; Indels 1; Gaps 1;

QY 38 AAPAAGGAGKAGEG-EIPAPLAGTYSKILVKGDTVKAGQTVLVEAMKETEINAP 96
DB 80 AAPASAPAAAAPAGAGPVPVAPLAGTITWKVLASGQTVAAAGEVLLILEAMKETEIRAA 139

QY 97 TDGKYEKVLKRDVAVGGGLIKI 121
DB 140 QAGTVRGIAVAKGDAVAVGDTLMTL 164

RESULT 16
US-08-485-607-106
Sequence 106, Application US/08485607
Patent No. 5792627
GENERAL INFORMATION:
APPLICANT: Robert Haselkorn and Piotr Gornicki
TITLE OF INVENTION: Cyanobacterial and Plant Acetyl-CoA
NUMBER OF SEQUENCES: 116
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: 321 No. 5792627th Clark Street
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60610
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy Disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII-DOS
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/485,607
FILING DATE: 07-JUN-1995
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/956,700
FILING DATE: 10/21/92
ATTORNEY/AGENT INFORMATION:
NAME: Thomas E. No. 5792627thrup
REGISTRATION NUMBER: 33,268
REFERENCE/DOCKET NUMBER: ARCD:058
TELECOMMUNICATION INFORMATION:
TELEPHONE: 1-312-744-0090
TELEFAX: 1-312-755-4489
INFORMATION FOR SEQ ID NO: 106:

SEQUENCE CHARACTERISTICS:
LENGTH: 165 amino acids
TYPE: Amino acid
STRANDEDNESS: Single
TOPOLOGY: Linear
MOLECULE TYPE: Peptide
US-08-485-607-106

Query Match 32.6%; Score 200.5; DB 1; Length 165;
Best Local Similarity 52.9%; Pred. No. 6.9e-14;
Matches 45; Conservative 10; Mismatches 29; Indels 1; Gaps 1;

QY 38 AAPAAGGAGKAGEG-EIPAPLAGTYSKILVKGDTVKAGQTVLVEAMKETEINAP 96
DB 80 AAPASAPAAAAPAGAGPVPVAPLAGTITWKVLASGQTVAAAGEVLLILEAMKETEIRAA 139

QY 97 TDGKYEKVLKRDVAVGGGLIKI 121
DB 140 QAGTVRGIAVAKGDAVAVGDTLMTL 164

RESULT 17
US-08-475-879-106
Sequence 106, Application US/08475879
Patent No. 5972644
Patent No. 5972644 5786170
GENERAL INFORMATION:
APPLICANT: Robert Haselkorn and Piotr Gornicki
TITLE OF INVENTION: Cyanobacterial and Plant Acetyl-CoA
NUMBER OF SEQUENCES: 116
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: 321 No. 5972644 5786170th Clark Street
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60610
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy Disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII-DOS
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/475,879
FILING DATE: 07-JUN-1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/956,700
FILING DATE: 10/21/92
ATTORNEY/AGENT INFORMATION:
NAME: Thomas E. No. 5972644 5786170thrup
REGISTRATION NUMBER: 33,268
REFERENCE/DOCKET NUMBER: ARCD:058
TELECOMMUNICATION INFORMATION:
TELEPHONE: 1-312-744-0090
TELEFAX: 1-312-755-4489
INFORMATION FOR SEQ ID NO: 106:
SEQUENCE CHARACTERISTICS:
LENGTH: 165 amino acids
TYPE: Amino acid
STRANDEDNESS: Single
TOPOLOGY: Linear
MOLECULE TYPE: Peptide
US-08-475-879-106

Query Match 32.6%; Score 200.5; DB 2; Length 165;
Best Local Similarity 52.9%; Pred. No. 6.9e-14;
Matches 45; Conservative 10; Mismatches 29; Indels 1; Gaps 1;

QY 38 AAPAAGGAGKAGEG-EIPAPLAGTYSKILVKGDTVKAGQTVLVEAMKETEINAP 96
DB 80 AAPASAPAAAAPAGAGPVPVAPLAGTITWKVLASGQTVAAAGEVLLILEAMKETEIRAA 139

```
QY      97 TDGKVEKLVKERDAVGGGGLIKI 121
      | | : | | | | | :
Db     140 QAGTVRGIAVKAGDAVAVGDTLMTL 164
```

```

RESULT 18
US-09-433-043B-106
: Sequence 106, Application US/09433043B
: Patent No. 6399342
: GENERAL INFORMATION:
: APPLICANT: HASELKORN, ROBERT
: APPLICANT: GORNICKI, PIORR
: TITLE OF INVENTION: CYANOBACTERIAL AND PLANT ACETYL-COA CARBOXYLASE
: FILE REFERENCE: ARCD:33805
: CURRENT APPLICATION NUMBER: US/09/433,043B
: CURRENT FILING DATE: 1999-10-25
: PRIOR APPLICATION NUMBER: 08/475,879
: PRIOR FILING DATE: 1995-06-07
: PRIOR APPLICATION NUMBER: 07/956,700
: PRIOR FILING DATE: 1992-10-02
: NUMBER OF SEQ ID NOS: 128
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 106
: LENGTH: 165
: TYPE: PRT
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: Description of Artificial Sequence: Synthetic
: OTHER INFORMATION: Peptide
: US-09-433-043B-106

```

Query Match	32.6%;	Score 200.5;	DB 4;	Length 165;
Best Local Similarity	52.9%;	Pred. No. 6.9e-14;		
Matches	45;	Conservative	10;	Mismatches 29;
				Indels 1;
				Gaps 1

[illegible]

```

RESULT 19
5252466-19
; Patent No. 5252466
; APPLICANT: CROMAN, JOHN E.
; TITLE OF INVENTION: FUSION PROTEINS HAVING A SITE FOR IN
; VIVO POST-TRANSLATION MODIFICATION AND METHODS OF MAKING AND
; PURIFYING THEM
; NUMBER OF SEQUENCES: 24
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/525,568
; FILING DATE: 18-MAY-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 354,266
; FILING DATE: 19-MAY-1989
; SEQ ID NO:19:
; LENGTH: 100
; 5252466-19

```

Query Match	31.5%	Score	193.5	DB	6	Length	100
Best Local Similarity	51.8%	Pred. No.	2e+13				
Matches	44	Conservative	10	Mismatches	30	Indels	1
						Gaps	1

[illegible]

Db 75 QAGTVRGIAVKAGDAVAVGDTLMTL 99

```

RESULT 20
US-09-220-081-2
: Sequence 2, Application US/09220081
: Patent No. 6171833
: GENERAL INFORMATION:
: APPLICANT: Sinskey, Anthony J.
: APPLICANT: Lessard, Philip A.
: APPLICANT: Willis, Laura B.
: APPLICANT: Stephanopoulos, Gregory
: TITLE OF INVENTION: Pyruvate Carboxylase from Corynebacterium glutamicum
: FILE REFERENCE: 1533.079000
: CURRENT APPLICATION NUMBER: US/09/220,081
: CURRENT FILING DATE: 1998-12-23
: NUMBER OF SEQ ID NOS: 36
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 2
: LENGTH: 1140
: TYPE: PRN
: ORGANISM: Corynebacterium glutamicum
: US-09-220-081-2

```

Query Match	25.98;	Score 159.5;	DB 4;	Length 1140;
Best Local Similarity	43.28;	Pred. No. 1.6e-08;		
Matches 32;	Conservative 15;	Mismatches 26;	Indels 1;	Gaps 1

QY	48	AGAGAEDEIPAPLACIVSKLVEGGIVACAGOTVLVLEMKMETELINPTDQKVERVLV	107
		: : : : : : : : : : : : : : : : : : : : : : : :	
Db	1067	ADSSNKGHVAAPFAGVYT-VTVAEGSEVKAAGDAVALIEMKMEATITTSYDGKIDRVYVP	1125
QY	108	ERDAVAGGCGGLIKI	121
		: : :	
Db	1126	AATKVEGGDLIVV	1139

```

RESULT 21
US-09-677-575-2
Sequence 2, Application US/09677575
Patent No. 6403351
GENERAL INFORMATION:
APPLICANT: Sinskey, Anthony J.
APPLICANT: Lessard, Philip A.
APPLICANT: Willis, Laura B.
APPLICANT: Stephanopoulos, Gregory
TITLE OF INVENTION: Pyruvate Carboxylase from Corynebacterium glutamicum
FILE REFERENCE: 1533.0790000
CURRENT APPLICATION NUMBER: US/09/677,575
CURRENT FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 09/220,081
PRIOR FILING DATE: 1998-12-23
NUMBER OF SEQ ID NOS: 36
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2
LENGTH: 1140
TYPE: prt
ORGANISM: Corynebacterium glutamicum
US-09-677-575-2

```

Query Match	25.98;	Score 159.5;	DB 4;	Length 1140;
Best Local Similarity	43.28;	Pred. No. 1.6e-08;		
Matches 32; Conservative	15;	Mismatches 26;	Indels 1;	Gaps . 1

QY	48	ACGACGAGGELPAPLPLAGVNSILVKEGTVYACGTVLVLEAMKETEINLPLOGKVEKVLV	107
		! : : : ! : : : ! : : : ! : : : ! : : : ! : : : ! : : : ! : : : ! : : : ! : : : !	
Db	1067	ADSSNKGHVAAPAGVLT-VTVAEGSEYKAGDAVALIEEMKKEATITTSVDSGKIDRVYVP	1125
QY	108	ERDAVVGCGGLIKI	121
		! : ! : : : ! : ! : : : ! : ! : : : ! : ! : : :	
Db	1126	AATKVEGGDLIVV	1139

RESULT 22
US-08-997-080-166
Sequence 166, Application US/08997080
Patent No. 5968524
GENERAL INFORMATION:
APPLICANT: WATSON, JAMES D.
APPLICANT: TAN, PAUL L.J.
TITLE OF INVENTION: METHODS AND COMPOUNDS FOR THE TREATMENT OF IMMUNOLOGICALLY-
NUMBER OF SEQUENCES: 194
CORRESPONDENCE ADDRESS:
ADDRESSEE: Law Offices of Ann W. Speckman
STREET: 2601 Elliott Avenue, Suite 4185
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98121
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/997,080
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
ATTORNEY/AGENT INFORMATION:
FILING DATE:
NAME: Sleath, Janet
REGISTRATION NUMBER: 37,007
REFERENCE/DOCKET NUMBER: 11000.1007
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-269-0565
TELEFAX: 206-269-0563
TELEX:
INFORMATION FOR SEQ ID NO: 166:
SEQUENCE CHARACTERISTICS:
LENGTH: 108 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-997-080-166
Query Match 24.4%; Score 150; DB 2; Length 108;
Best Local Similarity 46.5%; Pred. No. 8.5e-09;
Matches 33; Conservative 10; Mismatches 28; Indels 0; Gaps 0;
OY 52 GGEIPAPLAGTAVSKILYKEDGYKAGOTVLYLAKMKMETETINAPTDGKVEKVLKENDA 111
DB 33 GDERVDAPFASVWKVDVAVGDRVYAGPLALAKMKMETVLRAPADSVVTOILVSAGHL 92
OY 112 VOGGGLIKIG 122
DB 93 VDPGTPLYVG 103
RESULT 23
US-08-997-362-166
Sequence 166, Application US/08997362
Patent No. 5985287
GENERAL INFORMATION:
APPLICANT: Tan, Paul
APPLICANT: Hiyama, Jun
APPLICANT: Visser, Elizabeth
APPLICANT: Skinner, Margot
APPLICANT: Scott, Linda
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR
TITLE OF INVENTION: TREATMENT AND DIAGNOSIS OF MYCOBACTERIAL INFECTIONS
NUMBER OF SEQUENCES: 194
CORRESPONDENCE ADDRESS:
ADDRESSEE: Law Offices of Ann W. Speckman

STREET: 2601 Elliott Avenue, Suite 4185
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98121
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/997,362
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: U.S. Patent Application No. 5985287 08/873,970
FILING DATE: June 12, 1997
APPLICATION NUMBER: U.S. Patent Application No. 5985287 08/705,347
FILING DATE: August 29, 1996
ATTORNEY/AGENT INFORMATION:
NAME: Sleath, Janet
REGISTRATION NUMBER: 37,007
REFERENCE/DOCKET NUMBER: 11000.1002c2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-269-0565
TELEFAX: 206-269-0563
TELEX:
INFORMATION FOR SEQ ID NO: 166:
SEQUENCE CHARACTERISTICS:
LENGTH: 108 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-997-362-166
Query Match 24.4%; Score 150; DB 2; Length 108;
Best Local Similarity 46.5%; Pred. No. 8.5e-09;
Matches 33; Conservative 10; Mismatches 28; Indels 0; Gaps 0;
OY 52 GGEIPAPLAGTAVSKILYKEDGYKAGOTVLYLAKMKMETETINAPTDGKVEKVLKENDA 111
DB 33 GDERVDAPFASVWKVDVAVGDRVYAGPLALAKMKMETVLRAPADSVVTOILVSAGHL 92
OY 112 VOGGGLIKIG 122
DB 93 VDPGTPLYVG 103
RESULT 24
US-09-095-855-166
Sequence 166, Application US/09095855
Patent No. 6160093
GENERAL INFORMATION:
APPLICANT: Tan, Paul
APPLICANT: Visser, Elizabeth
APPLICANT: Skinner, Margot
APPLICANT: Prestidge, Ross
TITLE OF INVENTION: Compounds and Methods for
TITLE OF INVENTION: Treatment and Diagnosis of Mycobacterial Infections
NUMBER OF SEQUENCES: 208
CORRESPONDENCE ADDRESS:
ADDRESSEE: Law Offices of Ann W. Speckman
STREET: 2601 Elliott Avenue, Suite 4185
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98121
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 1, 2003, 07:50:32 ; Search time 8.38542 Seconds

(without alignments)
245.617 Million cell updates/sec

Title: US-09-987-485-2

Perfect score: 342
Sequence: 1 EGRIAPAPLAGVSKILVKEG.....KVLKERDAVGGGGLIKIG 70Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 200000000Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 100 summaries

Database :

Issued_Patents_AA:*
1: /cgn2_6/ptodata1/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata1/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata1/1aa/6A_COMB.pep:*
4: /cgn2_6/ptodata1/1aa/6B_COMB.pep:*
5: /cgn2_6/ptodata1/1aa/PCPMUS_COMB.pep:*
6: /cgn2_6/ptodata1/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	342	100.0	106	6	Patent No. 5252466
2	342	100.0	123	1	Sequence 107, App
3	342	100.0	123	1	Sequence 107, App
4	342	100.0	123	1	Sequence 107, App
5	342	100.0	123	1	Sequence 107, App
6	342	100.0	123	3	Sequence 107, App
7	342	100.0	123	3	Sequence 107, App
8	342	100.0	123	4	Sequence 107, App
9	342	100.0	123	4	Sequence 107, App
10	342	100.0	123	4	Sequence 107, App
11	336	98.2	123	4	Sequence 107, App
12	310	90.6	6	6	Patent No. 5252466
13	211	61.7	43	3	Sequence 127, App
14	167	48.8	170	4	Sequence 106, App
15	165	48.2	165	1	Sequence 106, App
16	165	48.2	165	1	Sequence 106, App
17	165	48.2	165	1	Sequence 106, App
18	165	48.2	165	1	Sequence 106, App
19	158	46.2	100	6	Patent No. 5252466
20	155.5	45.5	1140	4	Sequence 2, App1
21	155.5	45.5	1140	4	Sequence 2, App1
22	144	42.1	108	2	Sequence 166, App
23	144	42.1	108	2	Sequence 166, App
24	144	42.1	108	2	Sequence 166, App
25	144	42.1	108	2	Sequence 166, App
26	144	42.1	108	2	Sequence 166, App
27	144	42.1	243	4	Sequence 199, App

28	144	42.1	243	4	US-09-205-426-199	Sequence 199, App
29	138	40.4	1154	4	US-09-134-001C-3428	Sequence 3428, App
30	137	40.1	67	6	5252466-5	Patent No. 5252466
31	133	38.9	257	1	US-07-956-700B-105	Sequence 105, App
32	133	38.9	257	1	US-08-476-537-105	Sequence 105, App
33	133	38.9	257	1	US-08-485-607-105	Sequence 105, App
34	133	38.9	257	2	US-08-475-879-105	Sequence 105, App
35	133	38.9	257	4	US-09-433-043B-105	Sequence 105, App
36	133	38.9	694	4	US-09-433-043B-105	Sequence 105, App
37	131	38.3	75	6	5252466-3	Patent No. 5252466
38	126.5	37.0	182	2	US-08-422-560A-2	Sequence 2, App1
39	126.5	37.0	182	4	US-08-422-560A-2	Sequence 2, App1
40	126.5	37.0	182	4	US-08-422-560A-2	Sequence 2, App1
41	119.5	34.9	156	4	US-09-134-001C-3232	Sequence 3232, App
42	119.5	34.9	158	1	US-08-611-107-4	Sequence 4, App1
43	119.5	34.9	158	2	US-08-422-560A-4	Sequence 4, App1
44	119.5	34.9	158	4	US-08-422-560A-4	Sequence 4, App1
45	111	32.5	251	1	US-07-956-700B-94	Sequence 94, App1
46	111	32.5	251	1	US-08-476-537-94	Sequence 94, App1
47	111	32.5	251	1	US-08-485-607-94	Sequence 94, App1
48	111	32.5	251	2	US-08-475-879-94	Sequence 94, App1
49	111	32.5	251	4	US-09-433-043B-94	Sequence 94, App1
50	111	32.5	722	4	US-09-433-043B-125	Sequence 125, App
51	110	32.2	66	6	5252466-4	Patent No. 5252466
52	110	32.2	251	1	US-07-956-700B-84	Sequence 84, App1
53	110	32.2	251	1	US-08-476-537-84	Sequence 84, App1
54	110	32.2	251	1	US-08-485-607-84	Sequence 84, App1
55	110	32.2	251	2	US-08-475-879-84	Sequence 84, App1
56	110	32.2	251	4	US-09-433-043B-84	Sequence 84, App1
57	110	32.2	703	4	US-09-433-043B-124	Sequence 124, App
58	109	31.9	22	2	US-08-383-753-3	Sequence 3, App1
59	109	31.9	22	2	US-08-586-772-3	Sequence 3, App1
60	109	31.9	22	2	US-08-959-512-3	Sequence 3, App1
61	109	31.9	22	4	US-09-512-983-3	Sequence 3, App1
62	108	31.6	146	4	US-09-134-001C-3594	Sequence 3594, App
63	98.5	28.8	156	1	US-08-074-121-5	Sequence 5, App1
64	98.5	28.8	156	5	PCT-US94-06447-5	Sequence 5, App1
65	97.5	28.5	87	2	US-08-460-934-7	Sequence 7, App1
66	97.5	28.5	87	2	US-08-782-118-7	Sequence 7, App1
67	97.5	28.5	156	1	US-08-074-121-2	Sequence 2, App1
68	97.5	28.5	156	5	PCT-US94-06447-2	Sequence 2, App1
69	97.5	28.5	605	4	US-09-433-043B-123	Sequence 123, App
70	97.5	28.5	636	2	US-08-460-934-9	Sequence 9, App1
71	97.5	28.5	636	2	US-08-782-118-9	Sequence 9, App1
72	92	26.9	66	6	5252466-2	Patent No. 5252466
73	91	26.6	417	4	US-08-887-534A-38	Sequence 38, App1
74	91	26.6	417	4	US-08-887-534A-42	Sequence 42, App1
75	89	26.0	57	1	US-07-956-700B-74	Sequence 74, App1
76	89	26.0	57	1	US-08-476-537-74	Sequence 74, App1
77	89	26.0	57	1	US-08-485-607-74	Sequence 74, App1
78	89	26.0	57	2	US-08-475-879-74	Sequence 74, App1
79	89	26.0	57	2	US-08-475-879-74	Sequence 74, App1
80	88	25.7	435	4	US-09-134-001C-4622	Sequence 4622, App
81	88	25.7	435	4	US-09-134-001C-3418	Sequence 3418, App
82	87	25.4	437	4	US-07-754-018A-12	Sequence 12, App1
83	86.5	25.3	406	4	US-09-066-046-21	Sequence 21, App1
84	86.5	25.3	443	2	US-08-620-232A-3	Sequence 3, App1
85	85.5	25.0	443	2	US-09-005-232A-3	Sequence 3, App1
86	85.5	25.0	145	1	US-07-956-700B-111	Sequence 111, App
87	82.5	24.1	145	1	US-08-476-537-111	Sequence 111, App
88	82.5	24.1	145	1	US-08-485-607-111	Sequence 111, App
89	82.5	24.1	145	2	US-08-475-879-111	Sequence 111, App
90	82.5	24.1	145	2	US-08-475-879-111	Sequence 111, App
91	82.5	24.1	145	4	US-09-433-043B-111	Sequence 111, App
92	82.5	24.1	159	4	US-09-433-043B-122	Sequence 122, App
93	80	23.4	2089	1	US-08-418-833D-23	Sequence 23, App1
94	80	23.4	2089	1	US-08-418-833D-24	Sequence 24, App1
95	79	23.1	89	4	US-09-056-556-338	Sequence 238, App
96	79	23.1	89	4	US-09-072-596-233	Sequence 233, App
97	79	23.1	483	4	US-09-108-020-16	Sequence 16, App1
98	78	22.8	72	4	US-09-056-556-229	Sequence 229, App
99	78	22.8	72	4	US-09-072-596-224	Sequence 224, App
100	78	22.8	81	4	US-09-134-001C-3166	Sequence 3166, App

ALIGNMENTS

RESULT 1

5252466-6

Patent No. 5252466

APPLICANT: CROMAN, JOHN E.
TITLE OF INVENTION: FUSION PROTEINS HAVING A SITE FOR IN
VIVO POST-TRANSLATION MODIFICATION AND METHODS OF MAKING AND
PURIFYING THEM

NUMBER OF SEQUENCES: 24

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/525,568

FILING DATE: 18-MAY-1990

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 354,266

FILING DATE: 19-MAY-1989

SEQ ID NO: 6:

LENGTH: 106

5252466-6

Query Match 100.0%; Score 342; DB 6; Length 106;

Best Local Similarity 100.0%; Pred. No. 1.2e-38;

Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EGEIPAPLAGTVSKILVKEGDTYKAGQTVLVEAMKMEIINAPTDGKYEKVLKERDAY 60

DB 37 EGEIPAPLAGTVSKILVKEGDTYKAGQTVLVEAMKMEIINAPTDGKYEKVLKERDAY 96

QY 61 QGGGGLIKIG 70

DB 97 QGGGGLIKIG 106

RESULT 2

US-07-956-700B-107

Sequence 107, Application US/07956700B

Patent No. 5539092

GENERAL INFORMATION:

APPLICANT: Robert Haselkorn and Piotr Gornicki

TITLE OF INVENTION: Cyanobacterial and Plant Acetyl-CoA

NUMBER OF SEQUENCES: 116

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Arnold, White & Durkee

STREET: 321 No. 5539092th Clark Street

CITY: Chicago

STATE: Illinois

COUNTRY: USA

ZIP: 60610

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy Disk

COMPUTER: IBM PC Compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: ASCII-DOS

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/956,700B

FILING DATE: 19921002

CLASSIFICATION: 800

ATTORNEY/AGENT INFORMATION:

NAME: Thomas E. No. 5539092thrup

REGISTRATION NUMBER: 33,268

REFERENCE/DOCKET NUMBER: ARCD:058

TELECOMMUNICATION INFORMATION:

TELEPHONE: 1-312-744-0090

TELEFAX: 1-312-755-4489

INFORMATION FOR SEQ ID NO: 107:

SEQUENCE CHARACTERISTICS:

LENGTH: 123 amino acids

TYPE: Amino acid

STRANDEDNESS: Single

; TOPOLOGY: Linear

; MOLECULE TYPE: Peptide

US-07-956-700B-107

Query Match 100.0%; Score 342; DB 1; Length 123;

Best Local Similarity 100.0%; Pred. No. 1.4e-38;

Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EGEIPAPLAGTVSKILVKEGDTYKAGQTVLVEAMKMEIINAPTDGKYEKVLKERDAY 60

DB 54 EGEIPAPLAGTVSKILVKEGDTYKAGQTVLVEAMKMEIINAPTDGKYEKVLKERDAY 113

QY 61 QGGGGLIKIG 70

DB 114 QGGGGLIKIG 123

RESULT 3

US-08-476-537-107

Sequence 107, Application US/08476537

Patent No. 5756290th

GENERAL INFORMATION:

APPLICANT: Robert Haselkorn and Piotr Gornicki

TITLE OF INVENTION: Cyanobacterial and Plant Acetyl-CoA

NUMBER OF SEQUENCES: 116

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Arnold, White & Durkee

STREET: 321 No. 5756290th Clark Street

CITY: Chicago

STATE: Illinois

COUNTRY: USA

ZIP: 60610

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy Disk

COMPUTER: IBM PC Compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: ASCII-DOS

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/476,537

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/956,700

FILING DATE: 10/21/92

ATTORNEY/AGENT INFORMATION:

NAME: Thomas E. No. 5756290thrup

REGISTRATION NUMBER: 33,268

REFERENCE/DOCKET NUMBER: ARCD:058

TELECOMMUNICATION INFORMATION:

TELEPHONE: 1-312-744-0090

TELEFAX: 1-312-755-4489

INFORMATION FOR SEQ ID NO: 107:

SEQUENCE CHARACTERISTICS:

LENGTH: 123 amino acids

TYPE: Amino acid

STRANDEDNESS: Single

TOPOLOGY: Linear

MOLECULE TYPE: Peptide

Query Match 100.0%; Score 342; DB 1; Length 123;

Best Local Similarity 100.0%; Pred. No. 1.4e-38;

Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EGEIPAPLAGTVSKILVKEGDTYKAGQTVLVEAMKMEIINAPTDGKYEKVLKERDAY 60

DB 54 EGEIPAPLAGTVSKILVKEGDTYKAGQTVLVEAMKMEIINAPTDGKYEKVLKERDAY 113

QY 61 QGGGGLIKIG 70

DB 114 QGGGGLIKIG 123

RESULT 4
US-08-485-607-107
Sequence 107, Application US/08485607
Patent No. 5792627
GENERAL INFORMATION:
APPLICANT: Robert Haselkorn and Piotr Gornicki
TITLE OF INVENTION: Cyanobacterial and Plant Acetyl-CoA
NUMBER OF SEQUENCES: 116
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: 321 No. 5792627th Clark Street
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60610
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy Disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII-DOS
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/485,607
FILING DATE: 07-JUN-1995
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/956,700
FILING DATE: 10/21/92
ATTORNEY/AGENT INFORMATION:
NAME: Thomas E. No. 5792627thrup
REGISTRATION NUMBER: 33,268
REFERENCE/DOCKET NUMBER: ARCD:058
TELECOMMUNICATION INFORMATION:
TELEPHONE: 1-312-744-0090
TELEFAX: 1-312-755-4489
INFORMATION FOR SEQ ID NO: 107:
SEQUENCE CHARACTERISTICS:
LENGTH: 123 amino acids
TYPE: Amino acid
STRANDEDNESS: Single
TOPOLOGY: Linear
MOLECULE TYPE: Peptide
US-08-485-607-107
Query Match 100.0%; Score 342; DB 1; Length 123;
Best Local Similarity 100.0%; Pred. No. 1.4e-38;
Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 EGEIPAPLAGVSKILVKGDTVAGQTVLVLEAMKMETEINAPDGVVEVLYKERDVA 60
DB 54 EGEIPAPLAGVSKILVKGDTVAGQTVLVLEAMKMETEINAPDGVVEVLYKERDVA 113
QY 61 QGGGGLIKIG 70
DB 114 QGGGGLIKIG 123
RESULT 5
US-08-475-879-107
Sequence 107, Application US/08475879
Patent No. 5972644
Patent No. 5972644 5786170
GENERAL INFORMATION:
APPLICANT: Robert Haselkorn and Piotr Gornicki
TITLE OF INVENTION: Cyanobacterial and Plant Acetyl-CoA
NUMBER OF SEQUENCES: 116
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: 321 No. 5972644 5786170th Clark Street
CITY: Chicago
STATE: Illinois

COUNTRY: USA
ZIP: 60610
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy Disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII-DOS
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/475,879
FILING DATE: 07-JUN-1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/956,700
FILING DATE: 10/21/92
ATTORNEY/AGENT INFORMATION:
NAME: Thomas E. No. 5972644 5786170thrup
REGISTRATION NUMBER: 33,268
REFERENCE/DOCKET NUMBER: ARCD:058
TELECOMMUNICATION INFORMATION:
TELEPHONE: 1-312-744-0090
TELEFAX: 1-312-755-4489
INFORMATION FOR SEQ ID NO: 107:
SEQUENCE CHARACTERISTICS:
LENGTH: 123 amino acids
TYPE: Amino acid
STRANDEDNESS: Single
TOPOLOGY: Linear
MOLECULE TYPE: Peptide
US-08-475-879-107
Query Match 100.0%; Score 342; DB 2; Length 123;
Best Local Similarity 100.0%; Pred. No. 1.4e-38;
Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 EGEIPAPLAGVSKILVKGDTVAGQTVLVLEAMKMETEINAPDGVVEVLYKERDVA 60
DB 54 EGEIPAPLAGVSKILVKGDTVAGQTVLVLEAMKMETEINAPDGVVEVLYKERDVA 113
QY 61 QGGGGLIKIG 70
DB 114 QGGGGLIKIG 123
RESULT 6
US-07-687-819-1
Sequence 1, Application US/07687819
Patent No. 6072039
GENERAL INFORMATION:
APPLICANT: Haase, Ferdinand C.
APPLICANT: Haase, Ferdinand C.
TITLE OF INVENTION: Carrier Protein
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Rohm and Haas Company
STREET: Independence Mall West
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19105
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/687,819
FILING DATE: 19910520
CLASSIFICATION: 530
CLASSIFICATION: C07K13/00
CLASSIFICATION: C07K15/26
CLASSIFICATION: C12P21/00
CLASSIFICATION: 424/85.7
CLASSIFICATION: 530/351

CLASSIFICATION: 435/69.51
ATTORNEY/AGENT INFORMATION:
NAME: Draks, Jordan J.
REGISTRATION NUMBER: 22,029
REFERENCE/DOCKET NUMBER: 89-020
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-592-2478
TELEFAX: 215-592-2682
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 123 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Propionibacterium shermanii
PUBLICATION INFORMATION:
AUTHORS: Maloy, W L
AUTHORS: Bowen, B U
AUTHORS: Zwolinski, G K
AUTHORS: Kumar, K G
AUTHORS: Wood, H G
TITLE: Amino Acid Sequence of the Biotinyl Subunit
Patent No. 6072039
TITLE: from Transcarboxylase
JOURNAL: Journal of Biological Chemistry
VOLUME: 254
ISSUE: 22
PAGES: 11615-11622
DATE: No. 6072039 26-1979
RELEVANT RESIDUES IN SEQ ID NO: 1: FROM 1 TO 123
US-07-687-819-1

Query Match 100.0%; Score 342; DB 3; Length 123;
Best Local Similarity 100.0%; Pred. No. 1.4e-38;
Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EGEIPAPLAGTYSKIIYKESDVTYKAGQTVLVLEAMKMETEINAPTDGKYEKVLVKERDAY 60
DB 54 EGEIPAPLAGTYSKIIYKESDVTYKAGQTVLVLEAMKMETEINAPTDGKYEKVLVKERDAY 113
QY 61 QGGGGLIKIG 70
DB 114 QGGGGLIKIG 123

RESULT 7
US-09-433-043B-107
Sequence 107, Application US/09433043B
Patent No. 6399342
GENERAL INFORMATION:
APPLICANT: HASELKORN, ROBERT
APPLICANT: GORINCKI, PIOTR
TITLE OF INVENTION: CYANOBACTERIAL AND PLANT ACETYL-COA CARBOXYLASE
FILE REFERENCE: ARCD:338US
CURRENT APPLICATION NUMBER: US/09/433,043B
CURRENT FILING DATE: 1999-10-25
PRIOR APPLICATION NUMBER: 08/475,879
PRIOR FILING DATE: 1995-06-07
PRIOR APPLICATION NUMBER: 07/956,700
PRIOR FILING DATE: 1992-10-02
NUMBER OF SEQ ID NOS: 128
SOFTWARE: Patentln Ver. 2.1
SEQ ID NO 107
LENGTH: 123
TYPE: prt
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-433-043B-107

Query Match 100.0%; Score 342; DB 4; Length 123;

Best Local Similarity 100.0%; Pred. No. 1.4e-38;
Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EGEIPAPLAGTYSKIIYKESDVTYKAGQTVLVLEAMKMETEINAPTDGKYEKVLVKERDAY 60
DB 54 EGEIPAPLAGTYSKIIYKESDVTYKAGQTVLVLEAMKMETEINAPTDGKYEKVLVKERDAY 113
QY 61 QGGGGLIKIG 70
DB 114 QGGGGLIKIG 123

RESULT 8
US-09-364-083-2
Sequence 2, Application US/09364083
Patent No. 6210676
GENERAL INFORMATION:
APPLICANT: Callister, Steven M
APPLICANT: Lovrich, Steven D
APPLICANT: Scheil, Ronald F
APPLICANT: Jobe, Dean A
TITLE OF INVENTION: Compositions and Method Using the Borreliaacidal
TITLE OF INVENTION: Epitope(s) of Borrelia burgdorferi Outer Surface
TITLE OF INVENTION: Protein C (OspC) for the Diagnosis and Prevention of
FILE REFERENCE: Lyme Disease
CURRENT APPLICATION NUMBER: US/09/364,083
CURRENT FILING DATE: 1999-07-30
EARLIER APPLICATION NUMBER: 60/094,955
EARLIER FILING DATE: 1998-07-31
NUMBER OF SEQ ID NOS: 4
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO 2
LENGTH: 194
TYPE: prt
ORGANISM: Borrelia burgdorferi
US-09-364-083-2

Query Match 100.0%; Score 342; DB 4; Length 194;
Best Local Similarity 100.0%; Pred. No. 2.6e-38;
Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EGEIPAPLAGTYSKIIYKESDVTYKAGQTVLVLEAMKMETEINAPTDGKYEKVLVKERDAY 60
DB 53 EGEIPAPLAGTYSKIIYKESDVTYKAGQTVLVLEAMKMETEINAPTDGKYEKVLVKERDAY 112
QY 61 QGGGGLIKIG 70
DB 113 QGGGGLIKIG 122

RESULT 9
US-09-651-419-2
Sequence 2, Application US/09651419
Patent No. 6464985
GENERAL INFORMATION:
APPLICANT: Callister, Steven M
APPLICANT: Lovrich, Steven D
APPLICANT: Scheil, Ronald F
APPLICANT: Jobe, Dean A
TITLE OF INVENTION: Compositions and Method Using the Borreliaacidal
TITLE OF INVENTION: Epitope(s) of Borrelia burgdorferi Outer Surface
TITLE OF INVENTION: Protein C (OspC) for the Diagnosis and Prevention of
FILE REFERENCE: Lyme Disease
CURRENT APPLICATION NUMBER: US/09/651,419
CURRENT FILING DATE: 2000-08-30
PRIOR APPLICATION NUMBER: 60/094,955
PRIOR FILING DATE: 1998-07-31
NUMBER OF SEQ ID NOS: 4
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO 2
LENGTH: 194

;; TYPE: PRT
;; ORGANISM: Borrelia burgdorferi
US-09-651-419-2

Query Match
Best Local Similarity 100.0%; Score 342; DB 4; Length 194;
Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EGEIPAPLAGVSKILVKEGDTVAKGQTVLVLEAMKMETEINAPTDGKVEKVLKERDAV 60
DB 53 EGEIPAPLAGVSKILVKEGDTVAKGQTVLVLEAMKMETEINAPTDGKVEKVLKERDAV 112
QY 61 QGGGGLIKIG 70
DB 113 QGGGGLIKIG 122

RESULT 10
US-09-433-043B-128
; Sequence 128, Application US/09433043B
; Patent No. 6399342
; GENERAL INFORMATION:
; APPLICANT: HASELKORN, ROBERT
; APPLICANT: GORNICTI, PIOTR
; TITLE OF INVENTION: CYANOBACTERIAL AND PLANT ACETYL-COA CARBOXYLASE
; FILE REFERENCE: ARCD:33805
; CURRENT APPLICATION NUMBER: US/09/433, 043B
; CURRENT FILING DATE: 1999-10-25
; PRIOR APPLICATION NUMBER: 08/475, 879
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 07/956, 700
; PRIOR FILING DATE: 1992-10-02
; NUMBER OF SEQ ID NOS: 128
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 128
; LENGTH: 123
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Peptide
US-09-433-043B-128

Query Match
Best Local Similarity 98.2%; Score 336; DB 4; Length 123;
Matches 69; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EGEIPAPLAGVSKILVKEGDTVAKGQTVLVLEAMKMETEINAPTDGKVEKVLKERDAV 60
DB 54 EGEIPAPLAGVSKILVKEGDTVAKGQTVLVLEAMKMETEINAPTDGKVEKVLKERDAV 113
QY 61 QGGGGLIKIG 70
DB 114 QGGGGLIHIG 123

RESULT 11
5252466-1
; Patent No. 5252466
; APPLICANT: CROWAN, JOHN E.
; TITLE OF INVENTION: FUSION PROTEINS HAVING A SITE FOR IN
; VIVO POST-TRANSLATION MODIFICATION AND METHODS OF MAKING AND
; PURIFYING THEM
; NUMBER OF SEQUENCES: 24
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/525, 568
; FILING DATE: 18-MAY-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 354, 266
; FILING DATE: 19-MAY-1989
; SEQ ID NO: 1
; LENGTH: 66
5252466-1

Query Match
Best Local Similarity 90.6%; Score 310; DB 6; Length 66;
Matches 64; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 PAPIAGVSKILVKEGDTVAKGQTVLVLEAMKMETEINAPTDGKVEKVLKERDAV 64
DB 1 PAPIAGVSKILVKEGDTVAKGQTVLVLEAMKMETEINAPTDGKVEKVLKERDAV 60
QY 65 GLIKIG 70
DB 61 GLIKIG 66

RESULT 12
US-07-687-819-2
; Sequence 2, Application US/07687819
; Patent No. 6072039
; GENERAL INFORMATION:
; APPLICANT: Haase, Ferdinand C.
; APPLICANT: Cress, Dean E.
; TITLE OF INVENTION: Carrier Protein
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rohm and Haas Company
; STREET: Independence Mall West
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/687, 819
; FILING DATE: 19910520
; CLASSIFICATION: 530
; CLASSIFICATION: C07K13/00
; CLASSIFICATION: C07K15/26
; CLASSIFICATION: C12P21/00
; CLASSIFICATION: 424/85.7
; CLASSIFICATION: 530/351
; CLASSIFICATION: 435/69.51
; ATTORNEY/AGENT INFORMATION:
; NAME: Driks, Jordan J.
; REGISTRATION NUMBER: 22,029
; REFERENCE/DOCKET NUMBER: 89-020
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-592-2478
; TELEFAX: 215-592-2682
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 43 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
US-07-687-819-2

Query Match
Best Local Similarity 61.7%; Score 211; DB 3; Length 43;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 PAPIAGVSKILVKEGDTVAKGQTVLVLEAMKMETEINAPTDG 47
DB 1 PAPIAGVSKILVKEGDTVAKGQTVLVLEAMKMETEINAPTDG 43

RESULT 13
US-09-433-043B-127
; Sequence 127, Application US/09433043B

Patent No. 6399342
GENERAL INFORMATION:
APPLICANT: HASELKORN, ROBERT
TITLE OF INVENTION: CYANOBACTERIAL AND PLANT ACETYL-COA CARBOXYLASE
FILE REFERENCE: ARCD:338US
CURRENT APPLICATION NUMBER: US/09/433,043B
CURRENT FILING DATE: 1999-10-25
PRIOR APPLICATION NUMBER: 08/475,879
PRIOR FILING DATE: 1995-06-07
PRIOR APPLICATION NUMBER: 07/956,700
PRIOR FILING DATE: 1992-10-02
NUMBER OF SEQ ID NOS: 128
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 127
LENGTH: 170
TYPE: PRF
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-433-043B-127

Query Match 48.8%; Score 167; DB 4; Length 170;
Best Local Similarity 53.0%; Pred. No. 7.5e-15;
Matches 35; Conservative 9; Mismatches 22; Indels 0; Gaps 0;

QY 4 IPAPLAGTISKILVKEGDPVKAQGVTLVLEAKMKEIINAPFDGKVEKYLVERDAVGG 63
DB 104 VTAAPLAGTITWKVLASGQTVAAAGEVILLIEAKMKEIIRAAQAGTVRGIAVAGDAVAVG 163

QY 64 QGLIKI 69
DB 164 DTLMTL 169

RESULT 14
US-07-956-700B-106
Sequence 106, Application US/07956700B
Patent No. 5539092
GENERAL INFORMATION:
APPLICANT: Robert Haselkorn and Piotr Gornicki
TITLE OF INVENTION: Cyanobacterial and Plant Acetyl-CoA
NUMBER OF SEQUENCES: 116
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: 321 No. 5539092th Clark Street
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60610
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy Disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII-DOS
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/956,700B
FILING DATE: 19921002
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Thomas E. No. 5539092thrup
REGISTRATION NUMBER: 33,268
REFERENCE/DOCKET NUMBER: ARCD:058
TELECOMMUNICATION INFORMATION:
TELEPHONE: 1-312-744-0090
TELEFAX: 1-312-755-4489
INFORMATION FOR SEQ ID NO: 106:
SEQUENCE CHARACTERISTICS:
LENGTH: 165 amino acids
TYPE: Amino acid
STRANDEDNESS: Single

TOPOLOGY: Linear
MOLECULE TYPE: Peptide
US-07-956-700B-106

Query Match 48.2%; Score 165; DB 1; Length 165;
Best Local Similarity 53.0%; Pred. No. 1.3e-14;
Matches 35; Conservative 9; Mismatches 22; Indels 0; Gaps 0;

QY 4 IPAPLAGTISKILVKEGDPVKAQGVTLVLEAKMKEIINAPFDGKVEKYLVERDAVGG 63
DB 99 VTAAPLAGTITWKVLASGQTVAAAGEVILLIEAKMKEIIRAAQAGTVRGIAVAGDAVAVG 158

QY 64 QGLIKI 69
DB 159 DTLMTL 164

RESULT 15
US-08-476-537-106
Sequence 106, Application US/08476537
Patent No. 5756290
GENERAL INFORMATION:

APPLICANT: Robert Haselkorn and Piotr Gornicki
TITLE OF INVENTION: Cyanobacterial and Plant Acetyl-CoA
NUMBER OF SEQUENCES: 116
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: 321 No. 5756290th Clark Street
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60610
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy Disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII-DOS
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/476,537
FILING DATE:

CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/956,700
FILING DATE: 10/21/92
ATTORNEY/AGENT INFORMATION:
NAME: Thomas E. No. 5756290thrup
REGISTRATION NUMBER: 33,268
REFERENCE/DOCKET NUMBER: ARCD:058
TELECOMMUNICATION INFORMATION:
TELEPHONE: 1-312-744-0090
TELEFAX: 1-312-755-4489
INFORMATION FOR SEQ ID NO: 106:
SEQUENCE CHARACTERISTICS:
LENGTH: 165 amino acids
TYPE: Amino acid
STRANDEDNESS: Single
TOPOLOGY: Linear
MOLECULE TYPE: Peptide
US-08-476-537-106

Query Match 48.2%; Score 165; DB 1; Length 165;
Best Local Similarity 53.0%; Pred. No. 1.3e-14;
Matches 35; Conservative 9; Mismatches 22; Indels 0; Gaps 0;

QY 4 IPAPLAGTISKILVKEGDPVKAQGVTLVLEAKMKEIINAPFDGKVEKYLVERDAVGG 63
DB 99 VTAAPLAGTITWKVLASGQTVAAAGEVILLIEAKMKEIIRAAQAGTVRGIAVAGDAVAVG 158

QY 64 QGLIKI 69
DB 159 DTLMTL 164

RESULT 16
US-08-485-607-106
Sequence 106, Application US/08485607
Patent No. 5792627
GENERAL INFORMATION:
APPLICANT: Robert Haselkorn and Piotr Gornicki
TITLE OF INVENTION: Cyanobacterial and Plant Acetyl-CoA
NUMBER OF SEQUENCES: 116
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: 321 No. 5792627th Clark Street
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60610
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII-DOS
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/485,607
FILING DATE: 07-JUN-1995
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/956,700
FILING DATE: 10/21/92
ATTORNEY/AGENT INFORMATION:
NAME: Thomas E. No. 5792627thrup
REGISTRATION NUMBER: 33,268
REFERENCE/DOCKET NUMBER: ARCD:058
TELECOMMUNICATION INFORMATION:
TELEPHONE: 1-312-744-0090
FAX: 1-312-755-4489
INFORMATION FOR SEQ ID NO: 106:
SEQUENCE CHARACTERISTICS:
LENGTH: 165 amino acids
TYPE: Amino acid
STRANDEDNESS: Single
TOPOLOGY: Linear
MOLECULE TYPE: Peptide
US-08-485-607-106

Query Match 48.2%; Score 165; DB 1; Length 165;
Best Local Similarity 53.0%; Pred. No. 1.3e-14;
Matches 35; Conservative 9; Mismatches 22; Indels 0; Gaps 0;

QY 4 IPAPLAGTIVSKILVEGDTVAKGQTVLVLEAMKMETEINAPTDGKVEKVLVKEKRDVAVGG 63
DB 99 VTAPLAGTIVKVLASEGQTVAAAEVLLILEAMKMETEIRAAAGTIVGIAVKAGDAVAVG 158

QY 64 QGLIKI 69
DB 159 DTLMTL 164

RESULT 17
US-08-475-879-106
Sequence 106, Application US/08475879
Patent No. 5972644
Patent No. 5972644 5786170
GENERAL INFORMATION:
APPLICANT: Robert Haselkorn and Piotr Gornicki
TITLE OF INVENTION: Cyanobacterial and Plant Acetyl-CoA
NUMBER OF SEQUENCES: 116
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: 321 No. 5972644 5786170th Clark Street
CITY: Chicago
STATE: Illinois

COUNTRY: USA
ZIP: 60610
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII-DOS
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/475,879
FILING DATE: 07-JUN-1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/956,700
FILING DATE: 10/21/92
ATTORNEY/AGENT INFORMATION:
NAME: Thomas E. No. 5972644 5786170thrup
REGISTRATION NUMBER: 33,268
REFERENCE/DOCKET NUMBER: ARCD:058
TELECOMMUNICATION INFORMATION:
TELEPHONE: 1-312-744-0090
FAX: 1-312-755-4489
INFORMATION FOR SEQ ID NO: 106:
SEQUENCE CHARACTERISTICS:
LENGTH: 165 amino acids
TYPE: Amino acid
STRANDEDNESS: Single
TOPOLOGY: Linear
MOLECULE TYPE: Peptide
US-08-475-879-106

Query Match 48.2%; Score 165; DB 2; Length 165;
Best Local Similarity 53.0%; Pred. No. 1.3e-14;
Matches 35; Conservative 9; Mismatches 22; Indels 0; Gaps 0;

QY 4 IPAPLAGTIVSKILVEGDTVAKGQTVLVLEAMKMETEINAPTDGKVEKVLVKEKRDVAVGG 63
DB 99 VTAPLAGTIVKVLASEGQTVAAAEVLLILEAMKMETEIRAAAGTIVGIAVKAGDAVAVG 158

QY 64 QGLIKI 69
DB 159 DTLMTL 164

RESULT 18
US-09-433-043B-106
Sequence 106, Application US/09433043B
Patent No. 6399342
GENERAL INFORMATION:
APPLICANT: HASELKORN, ROBERT
TITLE OF INVENTION: CYANOBACTERIAL AND PLANT ACETYL-COA CARBOXYLASE
FILE REFERENCE: ARCD:338US
CURRENT APPLICATION NUMBER: US/09/433,043B
CURRENT FILING DATE: 1999-10-25
PRIOR APPLICATION NUMBER: 08/475,879
PRIOR FILING DATE: 1995-06-07
PRIOR APPLICATION NUMBER: 07/956,700
PRIOR FILING DATE: 1992-10-02
NUMBER OF SEQ ID NOS: 128
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 106
LENGTH: 165
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE: Description of Artificial Sequence: Synthetic
OTHER INFORMATION: Peptide
US-09-433-043B-106

Query Match 48.2%; Score 165; DB 4; Length 165;
Best Local Similarity 53.0%; Pred. No. 1.3e-14;
Matches 35; Conservative 9; Mismatches 22; Indels 0; Gaps 0;

INFORMATION FOR SEQ ID NO: 166:
SEQUENCE CHARACTERISTICS:
LENGTH: 108 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-997-080-166

Query Match 42.1%; Score 144; DB 2; Length 108;
Best Local Similarity 47.8%; Pred. No. 5e-12;
Matches 32; Conservative 9; Mismatches 26; Indels 0; Gaps 0;

QY 4 IPAPLAGTIVSKILYKGGDTVAGOTVLEAKMKETINAPTQDKVEKVLKERDAYOGG 63
DB 37 VDAPFASVWYVDVAVGDRVYVAGOPPLALEAKMKETVLRAPADGVVQIIVSAGHLVDPG 96
QY 64 OGLIKIG 70
DB 97 TPLVYVG 103

RESULT 23

US-08-997-362-166
Sequence 166, Application US/08997362
Patent No. 5985287

GENERAL INFORMATION:

APPLICANT: Tan, Paul
APPLICANT: Hiyyama, Jun
APPLICANT: Visser, Elizabeth
APPLICANT: Skinner, Margot
APPLICANT: Scott, Linda
APPLICANT: Prestidge, Ross
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR
TREATMENT AND DIAGNOSIS OF MYCOBACTERIAL INFECTIONS
NUMBER OF SEQUENCES: 194
CORRESPONDENCE ADDRESS:
ADDRESSEE: Law Offices of Ann W. Speckman
STREET: 2601 Elliott Avenue, Suite 4185
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98121
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/997,362
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: U.S. Patent Application No. 5985287 08/873,970
FILING DATE: June 12, 1997
APPLICATION NUMBER: U.S. Patent Application No. 5985287 08/705,347
FILING DATE: August 29, 1996
ATTORNEY/AGENT INFORMATION:
NAME: Sleath, Janet
REGISTRATION NUMBER: 37,007
REFERENCE/DOCKET NUMBER: 11000.1002c2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-269-0565
TELEFAX: 206-269-0563
TELEX:
INFORMATION FOR SEQ ID NO: 166:
SEQUENCE CHARACTERISTICS:
LENGTH: 108 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-997-362-166

Query Match 42.1%; Score 144; DB 2; Length 108;

Best Local Similarity 47.8%; Pred. No. 5e-12;
Matches 32; Conservative 9; Mismatches 26; Indels 0; Gaps 0;

QY 4 IPAPLAGTIVSKILYKGGDTVAGOTVLEAKMKETINAPTQDKVEKVLKERDAYOGG 63
DB 37 VDAPFASVWYVDVAVGDRVYVAGOPPLALEAKMKETVLRAPADGVVQIIVSAGHLVDPG 96
QY 64 OGLIKIG 70
DB 97 TPLVYVG 103

RESULT 24

US-09-095-855-166
Sequence 166, Application US/09095855
Patent No. 6160093

GENERAL INFORMATION:

APPLICANT: Tan, Paul
APPLICANT: Visser, Elizabeth
APPLICANT: Skinner, Margot
APPLICANT: Prestidge, Ross
TITLE OF INVENTION: Compounds and Methods for
Treatment and Diagnosis of Mycobacterial Infections
NUMBER OF SEQUENCES: 208
CORRESPONDENCE ADDRESS:
ADDRESSEE: Law Offices of Ann W. Speckman
STREET: 2601 Elliott Avenue, Suite 4185
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98121
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/095,855
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/705,347
FILING DATE: 29-AUG-1996
APPLICATION NUMBER: 08/873,970
FILING DATE: 12-JUN-1997
APPLICATION NUMBER: 08/997,362
FILING DATE: 23-DEC-1997
ATTORNEY/AGENT INFORMATION:
NAME: Sleath, Janet
REGISTRATION NUMBER: 37,007
REFERENCE/DOCKET NUMBER: 11000.1002c3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-269-0565
TELEFAX: 206-269-0563
TELEX:
INFORMATION FOR SEQ ID NO: 166:
SEQUENCE CHARACTERISTICS:
LENGTH: 108 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-095-855-166

US-09-095-855-166

Query Match 42.1%; Score 144; DB 4; Length 108;
Best Local Similarity 47.8%; Pred. No. 5e-12;
Matches 32; Conservative 9; Mismatches 26; Indels 0; Gaps 0;

QY 4 IPAPLAGTIVSKILYKGGDTVAGOTVLEAKMKETINAPTQDKVEKVLKERDAYOGG 63
DB 37 VDAPFASVWYVDVAVGDRVYVAGOPPLALEAKMKETVLRAPADGVVQIIVSAGHLVDPG 96
QY 64 OGLIKIG 70
DB 97 TPLVYVG 103

US-09-205-426-199
; Sequence 199, Application US/09205426
; Patent No. 6406704
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; TITLE OF INVENTION: Compounds and Methods for Treatment and
; TITLE OF INVENTION: Diagnosis of Mycobacterial Infections
; FILE REFERENCE: 11000.1002c4
; CURRENT APPLICATION NUMBER: US/09/205,426
; CURRENT FILING DATE: 1998-12-04
; EARLIER APPLICATION NUMBER: 09/095,855
; EARLIER FILING DATE: 1998-06-11
; EARLIER APPLICATION NUMBER: 08/997,362
; EARLIER FILING DATE: 1997-12-23
; EARLIER APPLICATION NUMBER: 08/873,970
; EARLIER FILING DATE: 1997-06-12
; EARLIER APPLICATION NUMBER: 08/705,347
; EARLIER FILING DATE: 1996-08-29
; NUMBER OF SEQ ID NOS: 208
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 199
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Mycobacterium vaccae
US-09-205-426-199

Query Match 42.1%; Score 144; DB 4; Length 243;
Best Local Similarity 47.8%; Pred. No. 1.5e-11;

Matches 32; Conservative 9; Mismatches 26; Indels 0; Gaps 0;

QY 4 IPAPLAGTYSKILVKEGDTYKAGQTVLVEAMKMETEINAPTDGKVEKVLKERDAVOGG 63
DB 172 VDAPEASVWVVDVAVGGRVAGQPLALEAMKMETEINAPTDGKVEKVLKERDAVOGG 231
QY 64 QGLIKIG 70
DB 232 TPLVYVG 238

RESULT 29
US-09-134-001C-3428
; Sequence 3428, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 3428
; LENGTH: 1154
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3428

Query Match 40.4%; Score 138; DB 4; Length 1154;
Best Local Similarity 43.9%; Pred. No. 7.8e-10;

Matches 29; Conservative 15; Mismatches 22; Indels 0; Gaps 0;

QY 4 IPAPLAGTYSKILVKEGDTYKAGQTVLVEAMKMETEINAPTDGKVEKVLKERDAVOGG 63
DB 1085 IGAQMPGVSVEYKVSVDVAVGQPLALEAMKMETEINAPTDGKVEKVLKERDAVOGG 1144
QY 64 QGLIKI 69
DB 1145 DLVEI 1150

RESULT 30
5252466-5
; Patent No. 5252466
; APPLICANT: CHONAN, JOHN E.
; TITLE OF INVENTION: FUSION PROTEINS HAVING A SITE FOR IN
; VIVO POST-TRANSLATION MODIFICATION AND METHODS OF MAKING AND
; PURIFYING THEM
; NUMBER OF SEQUENCES: 24
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/525,568
; FILING DATE: 18-MAY-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 354,266
; FILING DATE: 19-MAY-1989
; SEQ ID NO: 5
; LENGTH: 67
5252466-5

Query Match 40.1%; Score 137; DB 6; Length 67;
Best Local Similarity 50.0%; Pred. No. 2.3e-11;
Matches 29; Conservative 10; Mismatches 19; Indels 0; Gaps 0;

QY 6 APLAGTYSKILVKEGDTYKAGQTVLVEAMKMETEINAPTDGKVEKVLKERDAVOGG 63
DB 2 APMVGLVYKVLVKDGEKVEKQGPVLEAMKMEHVNKAPANGYVSGLEIKVGSVQDG 59

Search completed: May 1, 2003, 07:59:05
Job time: 9.38542 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4-p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 1, 2003, 07:49:48 ; Search time 18.4271 Seconds

(without alignments)
636.476 Million cell updates/sec

Title: US-09-987-485-1

Perfect score: 615
Sequence: 1 MKLKVTVNGTAYDVVDVDR.....KVLVKERDAVGGGGLIKIG 122Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : PIR_73: *
1: p1r1: *
2: p1r2: *
3: p1r3: *
4: p1r4: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	604.5	98.3	123	1 BKIP	biotin carboxyl ca
2	231.5	37.6	145	2 F75135	hypothetical prote
3	221	35.9	149	2 A71074	probable methylmal
4	205	33.3	129	2 D49094	methylnalonyl-CoA
5	200.5	32.6	596	2 A28088	oxaloacetate decar
6	198.5	32.3	134	2 T44984	methylnalonyl-CoA
7	190.5	31.0	655	2 B70432	pyruvate carboxyla
8	190	30.9	140	2 H69526	methylnalonyl-CoA
9	187.5	30.5	599	2 D81367	probable pyruvate
10	185	30.1	597	2 G82308	oxaloacetate decar
11	184	29.9	571	2 F71133	probable oxaloacet
12	176.5	28.7	576	2 H98247	biotin carboxylase
13	176.5	28.7	576	2 AC3038	biotin carboxylase
14	176	28.6	134	2 C72341	propionyl-CoA carb
15	176	28.6	134	2 T44608	pyruvate carboxyla
16	175.5	28.5	591	2 B44465	sodium ion pump ox
17	175.5	28.5	610	2 G84306	biotin carboxylase
18	173.5	28.2	186	2 D90418	hypothetical prote
19	173.5	28.2	591	2 AB0509	oxaloacetate decar
20	173.5	28.2	591	2 AE0909	oxaloacetate decar
21	173	28.1	567	2 F64453	oxaloacetate decar
22	170	27.6	1144	2 D97227	pyruvate carboxyla
23	168.5	27.4	607	2 F82966	probable transcarb
24	168.5	27.4	620	2 F70439	oxaloacetate decar
25	166	27.0	600	2 F70980	probable acCoA pro
26	165	26.8	1150	2 AB3978	pyruvate carboxyla
27	164	26.7	654	2 G87517	acetyl/propionyl-C
28	163	26.5	142	2 D69510	oxaloacetate decar
29	163	26.5	598	2 G86999	hypothetical prote

30	163	26.5	598	2 A55579	biotin carboxyl ca
31	162	26.3	591	2 S71009	biotin carboxylase
32	162	26.3	597	2 S71006	biotin carboxylase
33	161.5	26.3	1174	2 C97686	pyruvate carboxyla
34	161.5	26.3	1174	2 AE2911	pyruvate carboxyla
35	159	25.9	665	2 G97819	hypothetical prote
36	158.5	25.8	593	2 B71373	probable oxaloacet
37	157	25.5	161	2 A95049	hypothetical prote
38	157	25.5	161	2 G97919	acetyl-CoA carboxy
39	157	25.5	170	2 F87482	hypothetical prote
40	157	25.5	1146	2 AC1565	pyruvate carboxyla
41	157	25.5	1146	2 AH1208	pyruvate carboxyla
42	156	25.4	1150	2 G89881	pyruvate carboxyla
43	155	25.2	167	2 T44291	biotin carboxyl ca
44	153.5	25.0	436	2 A53568	methylnalonyl-Co
45	152.5	24.8	162	2 A97521	biotin carboxyl ca
46	152.5	24.8	162	2 AC2740	hypothetical prote
47	152	24.7	655	2 A83395	probable acyl-CoA
48	150	24.4	70	2 E83791	hypothetical prote
49	150	24.4	155	2 H86721	hypothetical prote
50	150	24.4	665	2 C71667	propionyl-CoA carb
51	147	23.9	1078	2 D87647	hypothetical prote
52	146.5	23.8	1127	2 D70671	pyruvate carboxyla
53	145	23.6	156	2 AH3384	biotin carboxyl ca
54	144	23.4	661	2 F83284	probable biotin ca
55	144	23.4	1185	2 T39734	pyruvate carboxyla
56	143.5	23.3	1137	2 E86708	pyruvate carboxyla
57	143	23.3	730	2 G86161	hypothetical prote
58	141.5	23.0	678	2 AG3492	propionyl-CoA carb
59	141	22.9	568	2 C69014	oxaloacetate decar
60	139	22.6	187	2 C75558	acetyl-CoA carboxy
61	138.5	22.5	157	2 S73204	acetyl-CoA carboxy
62	138	22.4	1180	2 S46094	pyruvate carboxyla
63	136.5	22.2	158	2 T30279	biotin carboxyl ca
64	133.5	21.7	159	2 D97338	biotin carboxyl ca
65	133.5	21.7	163	2 C81708	acetyl-coenzyme A
66	133	21.6	148	2 F69685	pyruvate carboxyla
67	133	21.6	1178	1 QYBYP	pyruvate carboxyla
68	131	21.3	164	2 G71553	probable biotin ca
69	131	21.3	1207	2 H87475	urea amidolyase-re
70	130	21.1	630	1 XXECP	dihydrolipoamide S
71	130	21.1	665	2 T02793	propionyl CoA carb
72	130	21.1	731	2 T06360	methylnalonyl-Co
73	130	21.1	1158	2 AE3285	pyruvate carboxyla
74	129.5	21.1	182	2 B53311	acetyl-CoA carboxy
75	129.5	21.1	182	2 A12437	biotin carboxyl ca
76	129	21.0	547	2 H83018	dihydrolipoamide a
77	129	21.0	667	2 C70719	biotin carboxylase
78	129	21.0	1178	1 A47255	pyruvate carboxyla
79	128	20.8	544	2 C42653	dihydrolipoamide S
80	128	20.8	688	2 T21641	hypothetical prote
81	127	20.7	590	2 T42202	probable acyl-CoA
82	127	20.7	590	2 T42206	probable acyl-CoA
83	126	20.5	630	2 G85494	hypothetical prote
84	126	20.5	630	2 G90643	hypothetical prote
85	126	20.5	662	2 G95929	probable methylcro
86	125.5	20.4	654	2 B55571	biotin carboxyl ca
87	124	20.2	196	2 G82341	acetyl-CoA carboxy
88	123.5	20.1	638	1 XXAV	dihydrolipoamide S
89	122.5	19.9	629	2 AH0521	dihydrolipoamide S
90	122	19.8	553	2 B55514	acetyl-CoA carboxy
91	121.5	19.8	151	2 E81033	biotin carboxyl ca
92	121.5	19.8	154	2 D70418	hypothetical prote
93	121.5	19.7	154	2 G89932	hypothetical prote
94	121	19.7	1178	1 J02460	pyruvate carboxyla
95	120.5	19.6	1095	2 B83471	probable pyruvate
96	120	19.5	724	2 T16187	hypothetical prote
97	119	19.3	590	1 T35297	probable dihydrol
98	118.5	19.3	704	2 A34337	pyruvate carboxyla
99	118.5	19.3	1195	2 T43735	propionyl-CoA carb
100	118.5	19.3	1195	2 T43735	pyruvate carboxyla

ALIGNMENTS

RESULT 1

B:IP

biotin carboxyl carrier protein [validated] - Propionibacterium freudenreichii subsp. sh
N:Altenate names: methylmalonyl-CoA carboxyltransferase biotin carboxyl carrier protein
C:Species: Propionibacterium freudenreichii subsp. shermanii
C:Date: 31-Mar-1980 #sequence_revision 31-Mar-1980 #text_change 01-Feb-2002
C:Accession: A03401

R:Maloy, W.L.; Bowlen, B.O.; Zwolinski, G.K.; Kumar, K.G.; Wood, H.G.; Ericsson, L.H.; W
J. Biol. Chem. 254, 11615-11622, 1979
A:Title: Amino acid sequence of the biotinyl subunit from transcarboxylase.
A:Reference number: A03401; MUID:80049796; PMID:40985

A:Accession: A03401
A:Molecule type: Protein
A:Residues: 1-123 <MAL>

C:Comment: Six or 12 chains of biotin carboxyl carrier protein (BCCP) are found in the t
onyl coenzyme A to BCCP and (2) from BCCP to pyruvate, forming oxalacetate.
C:Superfamily: biotin carboxyl carrier protein; lipoyl/biotin-binding homology

C:Keywords: biotin binding
F:50-123/Domain: lipoyl/biotin-binding homology <LPB>
F:89/Binding site: biotin (lys) (covalent) #status experimental

Query Match 98.3%; Score 604.5; DB 1; Length 123;
Best Local Similarity 99.2%; Pred. No. 2.9e-41;
Matches 122; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 MKLKVNGTAYDVVDVDSHNPMTILFGGTGAPAP-AAAGAGAGAGAGGELIPAP 59
Db 1 MKLKVNGTAYDVVDVDSHNPMTILFGGTGAPAPRAAGAGAGAGAGGELIPAP 60

QY 60 LACTVSKILKEGDTYKAGTIVLLEAMKMETENAPTDCKVEKVLKEDAVOGGGLI 119
Db 61 LACTVSKILKEGDTYKAGTIVLLEAMKMETENAPTDCKVEKVLKEDAVOGGGLI 120

QY 120 KIG 122
Db 121 KIG 123

RESULT 2

F75135

methylmalonyl-CoA decarboxylase gamma chain PAB1771 - Pyrococcus abyssi (strain Orsay)

C:Species: Pyrococcus abyssi
C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Jun-2000
C:Accession: F75135

R:anonymous, Genoscope
submitted to the EMBL Data Library, July 1999

A:Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome stru
A:Reference number: A75001
A:Accession: F75135

A:Status: preliminary
A:Molecule type: DNA

A:Residues: 1-145 <KAW>
A:Cross-references: GB:AJ248285; GB:AL096836; NID:g5458067; PIDN:CAB49799.1; PID:g545831

A:Experimental source: strain Orsay
C:Genetics:

A:Gene: PAB1771

C:Superfamily: biotin carboxyl carrier protein; lipoyl/biotin-binding homology

Query Match 37.6%; Score 231.5; DB 2; Length 145;
Best Local Similarity 41.4%; Pred. No. 1.2e-11;
Matches 60; Conservative 18; Mismatches 44; Indels 23; Gaps 3;

QY 1 MKLKVNGTAYDVVDVDSHNPMTILFGGTGAPAP-AAAGAGAGAGAGGELIPAP 40
Db 1 MKKVYVNGKEVEVEEVPKFRVTLGKTYEVANIGIQVAPQTVATPAPPTPT 60

QY 41 PA---AGGAGAGAGAGGELIPAPAGTIVSKILVEGDTYKAGTIVLLEAMKMETENAP 97
Db 61 PPPVAPPTPPVQASENVVATAPMPKVKILVQEGQVKGGLITLLEAMKMETENAP 120

QY 98 DGKVEKVLKERDAVGGGGLIKIG 122
Db 121 DGKVEKVLKERDAVGGGGLIKIG 145

RESULT 3

A71074

probable methylmalonyl-CoA decarboxylase gamma chain - Pyrococcus horikoshii

C:Species: Pyrococcus horikoshii
C:Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 20-Jun-2000
C:Accession: A71074

R:Kawarayashi, Y.; Sawada, M.; Horikawa, H.; Halkawa, Y.; Hino, Y.; Yamamoto, S.; Se
M.; Ohtuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kusida, N.; Ogu
DNA Res. 5, 55-76, 1998

A:Title: Complete sequence and gene organization of the genome of a hyper-thermophil
A:Reference number: A71000; MUID:98344137; PMID:9679194

A:Accession: A71074
A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA
A:Residues: 1-149 <KAW>

A:Cross-references: GB:AP000005; NID:g3236132; PIDN:BAA30387.1; PID:g3257704
A:Experimental source: strain OT3

A:Note: this accession replaces an interim accession for a sequence replaced by Genba
C:Genetics:

A:Gene: PH1284

C:Superfamily: biotin carboxyl carrier protein; lipoyl/biotin-binding homology
F:76-149/Domain: lipoyl/biotin-binding homology <LPB>

Query Match 35.9%; Score 221; DB 2; Length 149;
Best Local Similarity 36.3%; Pred. No. 8.6e-11;
Matches 53; Conservative 21; Mismatches 48; Indels 24; Gaps 1;

QY 1 MKLKVNGTAYDVVDVDSHNPMTILFGGTGAPAP-AAAGAGAGAGAGGELIPAP 59
Db 4 MKKVYVNGKEVEVEEVPKFRVTLGKTYEVANIGIQVAPQTVATPAPPTPT 63

QY 37 GAPAPAAAGAGAGAGGELIPAPAGTIVSKILVEGDTYKAGTIVLLEAMKMETENAP 96
Db 64 PPTPAPAPSSKTVSENVVATAPMPKVKILVAVGDRVAVGGGLITLLEAMKMETENAP 123

QY 97 TDGKVEKVLKERDAVGGGGLIKIG 122
Db 124 RDGVKRLVKEGEAVDTGPTLIEIG 149

RESULT 4

D49094

methylmalonyl-CoA decarboxylase (EC 4.1.1.41) gamma chain - Veillonella parvula

C:Species: Veillonella parvula
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 20-Aug-1999
C:Accession: D49094

R:Huder, J.B.; Dimroth, P.
J. Biol. Chem. 268, 24564-24571, 1993

A:Title: Sequence of the sodium ion pump methylmalonyl-CoA decarboxylase from Veillon
A:Reference number: A49094; MUID:94043308; PMID:8227015

A:Accession: D49094
A:Status: preliminary

A:Molecule type: DNA
A:Residues: 1-129 <HND>

A:Cross-references: GB:L22208; NID:g415592; PIDN:AC36823.1; PID:g415596
C:Superfamily: biotin carboxyl carrier protein; lipoyl/biotin-binding homology

C:Keywords: carbon-carbon lyase; carboxy-lyase
F:57-129/Domain: lipoyl/biotin-binding homology <LPB>

Query Match 33.3%; Score 205; DB 2; Length 129;
Best Local Similarity 39.4%; Pred. No. 1.4e-09;
Matches 50; Conservative 15; Mismatches 36; Indels 6; Gaps 2;

QY 2 KUKVTYNGTAYDVVDVDSHNPMTILFGGTGAPAPAAAGAGAG---KAGEGE 55
Db 3 KFNVTYNGTAYDVVEVVEVAAAPAPAPAPAPAPAPAPAPAPAPAPAPAPAP 62

OY 56 IPAPLAVTSKILYKEDGTAKAGTIVYLEAMKMETEINAPTDCKVEKYLKERDAVOGG 115
 DB 63 VKAPMPKILTSVAASAGAVKGGTLLILEAMKQNEIAPHPDAVSEVRSANQOTVSTG 122
 OY 116 OGLIKIG 122
 DB 123 DDMVYLG 129

RESULT 5

A28088
 oxaloacetate decarboxylase (EC 4.1.1.3) alpha chain - Klebsiella pneumoniae
 N:Alternate names: oxalate beta-decarboxylase, alpha chain
 C:Species: Klebsiella pneumoniae
 C:Date: 19-May-1989 #sequence_revision 19-May-1989 #text_change 11-Jan-2002
 C:Accession: A28088
 R:Schwarz, E.; Oesterhelt, D.; Reinke, H.; Beyreuther, K.; Dimroth, P.
 J. Biol. Chem. 263, 9640-9645, 1988
 A:Title: The sodium ion translocating oxaloacetate decarboxylase of Klebsiella pneumoniae
 A:Reference number: A28088; MUID:88257085; PMID:2454915
 A:Accession: A28088
 A:Molecule type: DNA
 A:Residues: 1-596 <SCCH>
 A:Cross-references: EMBL:J03885; NID:g149288; PIDN:AA25120.1; PID:g149289
 C:Superfamily: Klebsiella pneumoniae oxaloacetate decarboxylase alpha chain; lipoyl/biotin
 C:Keywords: biotin binding; carbon-carbon lyase; carboxy-lyase; sodium pump
 F:523-596/Domain: lipoyl/biotin-binding homology <LBPB>
 F:562/Binding site: biotin (Lys) (covalent) #status predicted

Query Match 32.6%; Score 200.5; DB 2; Length 596;
 Best Local Similarity 52.9%; Pred. No. 1.5e-08;
 Matches 45; Conservative 10; Mismatches 29; Indels 1; Gaps 1;

OY 38 APAPAGACAGKAGEG-ELPAPLAVTSKILYKEDGTAKAGTIVYLEAMKMETEINAP 96
 DB 511 APAPASAPAAAPAGAGTPTVAPLAGITIKVLASEGQTVAGVEVLILEAMKMETEIRAA 570
 OY 97 TDGVEVYLVKERDAVOGGGLIKI 121
 DB 571 QAGTVRGIAVAKGDAVAVGDTLMTL 595

RESULT 6

T44984
 methylmalonyl-CoA decarboxylase (EC 4.1.1.41) gamma chain [imported] - Propionigenium m
 C:Species: Propionigenium modestum
 C:Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 21-Jul-2000
 C:Accession: T44984
 R:Botl, M.; Pfister, K.; Burda, P.; Kalbermatter, O.; Woehlke, G.; Dimroth, P.
 Eur. J. Biochem. 250, 590-599, 1997
 A:Title: Methylmalonyl-CoA decarboxylase from Propionigenium modestum: Cloning and sequ
 A:Reference number: 222888; MUID:98088990; PMID:9428714
 A:Accession: T44984
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-134 <BOT>
 A:Cross-references: EMBL:AV002015; NID:g2706397; PIDN:CAAO5139.1; PID:g2706400
 A:Experimental source: DSM 2376
 C:Comment: Methylmalonyl-CoA decarboxylase catalyses the only energy-conserving step dur
 1-CoA is coupled to the vectorial transport of Na⁺ across the cytoplasmic membrane, the
 C:Genetics:
 A:Note: mmcd
 C:Complex: heterotetramer [validated, MUID:98088990]
 C:Function:
 A:Description: EC 4.1.1.41 [validated, MUID:98088990]
 A:Note: specific activity up to 25 U/mg protein; Km value for (S)-methylmalonyl-CoA of a
 C:Superfamily: biotin carboxyl carrier protein; lipoyl/biotin-binding homology
 C:Keywords: carbon-carbon lyase; carboxy-lyase

Query Match 32.3%; Score 198.5; DB 2; Length 134;
 Best Local Similarity 38.5%; Pred. No. 4.7e-09;
 Matches 52; Conservative 12; Mismatches 44; Indels 27; Gaps 3;

OY 4 KTVNGTAVDVVDVDSKSHENPMGTILFGGCTG-----APAPAG 44
 DB 5 KTVNGTEYDAV-----EMGAAVVASAPAPAPAPAPAPAPAPAPAPAPAPAPAKT 57
 OY 45 GAGGAKGEGEIPAPLAVTSKILYKEDGTAKAGTIVYLEAMKMETEINAPTDCKVEK 104
 DB 58 AAGAG-AGANTVAPMPGTILLVNGCHAGDKVSKGDTLVLEAMKMETEINAPHDGVSEV 116
 OY 105 LVKERDAVOGGGLI 119
 DB 117 RVQOGASVNMADILV 131

RESULT 7

B70432
 pyruvate carboxylase c-terminal domain - Aquifex aeolicus
 C:Species: Aquifex aeolicus
 C:Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 18-Aug-2000
 C:Accession: B70432
 R:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.;
 V.

Nature 392, 353-358, 1998
 A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
 A:Reference number: A70300; MUID:98196666; PMID:9537320
 A:Accession: B70432
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-655 <AQF>
 A:Cross-references: GB:AE000744; NID:g2983891; PIDN:AA07445.1; PID:g2983894; GB:AE00
 A:Experimental source: strain VFS
 C:Genetics:
 A:Gene: PYCA
 C:Superfamily: Klebsiella pneumoniae oxaloacetate decarboxylase alpha chain; lipoyl/b
 F:572-646/Domain: lipoyl/biotin-binding homology <LBPB>
 F:612/Binding site: biotin (Lys) (covalent) #status predicted

Query Match 31.0%; Score 190.5; DB 2; Length 655;
 Best Local Similarity 36.8%; Pred. No. 9.9e-08;
 Matches 49; Conservative 19; Mismatches 52; Indels 13; Gaps 3;

OY 2 KLVTVNGTAVDV-----VVDKSHENPMGTILFGGCTGAPAPAGAGAG 49
 DB 513 KFKVRIEVSVDASGKPRKYVRINNRLEIQLKPREKALPQGGAGTQVSAEEEGIP 572
 OY 50 KAGE-GEIPAPLAVTSKILYKEDGTAKAGTIVYLEAMKMETEINAPTDCKVEKYLKE 108
 DB 573 KATRPDVTTPMPKPVKYLKKEBPVQGGTVAIVLEAMKMETEINAPIDQIVKIRAKP 632
 OY 109 RDAVOGGGGLIKI 121
 DB 633 GDQVNPQDAIMRI 645

RESULT 8

H69526
 methylmalonyl-CoA decarboxylase, biotin carboxyl carrier subunit (mmcd) homolog - Arc
 C:Species: Archaeoglobus fulgidus
 C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 20-Aug-1999
 C:Accession: H69526
 R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dod
 ; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E
 Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
 Nature 390, 364-370, 1997
 A:Authors: Overbeek, T.; Cotton, M.D.; Spriggs, T.; Artach, P.; Kahne, B.P.; Sykes,
 Smith, H.O.; Woese, C.R.; Venter, J.C.
 A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing arch
 A:Reference number: A69250; MUID:98049345; PMID:9389475
 A:Accession: H69526
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-140 <RLE>
 A:Cross-references: GB:AE000952; GB:AE000782; NID:g2669275; PIDN:AA889036.1; PID:g264
 C:Superfamily: biotin carboxyl carrier protein; lipoyl/biotin-binding homology

A; RESIDUES: 1-39 / <HEL>

Cross-references: GB:AE007870: PIDN:AAK89506.1: PTD:q15159380: GSPDB:GN00170

52

A:Residues: 1-591 <MOE>
 A:Cross-references: GB:M96434; NID:g154195; PIDN:AAA02973.1; PID:g408892
 A:Experimental source: LT2
 A:Note: sequence extracted from NCBI backbone
 C:Superfamily: Klebsiella pneumoniae oxalacetate decarboxylase alpha chain; lipoyl/biotin
 F:518-591/Domain: lipoyl/biotin-binding homology <LPB>
 F:557/Binding site: biotin (Lys) (covalent) #status predicted

Query Match 28.5%; Score 175.5; DB 2; Length 591;
 Best Local Similarity 37.5%; Pred. No. 1.4e-06;
 Matches 48; Conservative 10; Mismatches 41; Indels 29; Gaps 3;

QY 5 VTVNGTAYDVVDVVDKSHENPMGTILFGG-----TGGAAPAPAGAGAGAGAG 53
 DB 481 VEVEGKAFVYKVS-----DGGDISQLTAAPVPAASAPVQAAPAGAGT--- 523
 QY 54 GEIPAPLAGTAVSKILVEKGTAVKAGOTVLEAMKMETETINAPDGEVYVYKRDVAVQ 113
 DB 524 -PVTAPLAGINIKVIAIEGQTVAGSDVLLILEAMKMETETIRAAQAGTVRGIAVSGDAVS 582
 QY 114 GGQGLIKI 121
 DB 583 VGDITMTL 590

RESULT 17
 G84306
 biotin carboxylase [imported] - Halobacterium sp. NRC-1
 C:Species: Halobacterium sp. NRC-1
 C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001
 C:Accession: G84306

R:19; M.V.; Kennedy, S.P.; Mahaitas, G.G.; Bergquist, B.; Pan, M.; Shukla, H.D.; Laaky, S.; Leltlauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Javid Jung, K.R.; Alam, M.; Freitas, T.
 Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
 A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ehardt, H.; Lowe, T.M.; Li
 A:Title: Genome sequence of Halobacterium species NRC-1.
 A:Reference number: AB4160; MID:20504483; PMID:11016950
 A:Accession: G84306
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-610 <STO>
 A:Cross-references: GB:AE004437; NID:g10581019; PIDN:AA019819.1; GSPDB:GN00138
 C:Genetics:
 A:Gene: acc
 C:Superfamily: propionyl-CoA carboxylase alpha chain; biotin carboxylase homology; lipoyl/b

Query Match 28.5%; Score 175.5; DB 2; Length 610;
 Best Local Similarity 38.1%; Pred. No. 1.4e-06;
 Matches 48; Conservative 18; Mismatches 43; Indels 17; Gaps 5;

QY 5 VTVNGTAYDVVDVVDKSHENPMGTILFGG-----GTGAPAPAGAGAGAGAG--GEGE- 55
 DB 492 VEVEGKAFVYKVS-----ERGAQFAPEADTGGGPREP-AGGADGETVEGDEGT 543
 QY 56 IPAPLAGTAVSKILVEKGTAVKAGOTVLEAMKMETETINAPDGEVYVYKRDVAVQ 115
 DB 544 VTLEMOGTILDVAVSBDADAGDVLEAMKMETETIRAAQAGTVRGIAVSGDVSMD 603
 QY 116 GQGLIKI 121
 DB 604 DVLVVI 609

RESULT 18
 D90418
 hypothetical protein SS02464 [imported] - Sulfolobus solfataricus
 C:Species: Sulfolobus solfataricus
 C:Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 24-May-2001
 C:Accession: D90418
 R:She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan
 Jong, I.; Jeffries, A.C.; Koser, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, F
 arett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.

submitted to GenBank, April 2001
 A:Description: Sulfolobus solfataricus complete genome.
 A:Reference number: A99139
 A:Accession: D90418

A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-186 <NR>
 A:Cross-references: GB:AE00641; NID:g13815767; PIDN:AA442603.1; GSPDB:GN00155
 C:Genetics:
 A:Gene: SS02464

Query Match 28.2%; Score 173.5; DB 2; Length 186;
 Best Local Similarity 40.0%; Pred. No. 6.2e-07;
 Matches 50; Conservative 14; Mismatches 42; Indels 19; Gaps 4;

QY 7 VNGTAYDVVDVVDKSHENPMGTILFGG-----TGGAAPAPAGAGAGAGAGAG 56
 DB 70 INGGKTYVIESD-----GTILFNHDFLRDKYETPIKEERVEETIRCK--EGET 120
 QY 57 PAPLAGTAVSKILVEKGTAVKAGOTVLEAMKMETETINAPDGEVYVYKRDVAVQ 116
 DB 121 VSPILGFRVYKIRVEKGDVAVKQGPLLSIEMKAEYVSISSIGIVOKILKEGQVKKD 180
 QY 117 GLIKI 121
 DB 181 ILIYI 185

RESULT 19
 AB0509

oxalacetate decarboxylase (EC 4.1.1.3) alpha chain [imported] - Salmonella enterica
 C:Species: Salmonella enterica subsp. enterica serovar Typhi
 A:Note: this species has also been called Salmonella typhi
 C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 27-Nov-2001
 C:Accession: AB0509
 R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Church
 Th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Fair
 Nature 413, 848-852, 2001
 A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens,
 A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica se
 A:Reference number: AB0509; PMID:11677608
 A:Accession: AB0509
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-591 <PAR>
 A:Cross-references: GB:AL513382; PIDN:CAD01210.1; PID:g16501339; GSPDB:GN00176
 C:Genetics:
 A:Gene: cadA
 C:Superfamily: Klebsiella pneumoniae oxalacetate decarboxylase alpha chain; lipoyl/b

Query Match 28.2%; Score 173.5; DB 2; Length 591;
 Best Local Similarity 40.2%; Pred. No. 2e-06;
 Matches 47; Conservative 14; Mismatches 49; Indels 7; Gaps 3;

QY 5 VTVNGTAYDVVDVVDKSHENPMGTILFGGCTGAPAPAGAGAGAGAGAGAGAGAGAG 64
 DB 481 VEVEGKAFVYKVS-DGGDISQLTTAV--PAASAPVQAAPAGAGT---PVTAPLAGNI 533
 QY 65 SKILVEGTVKAGOTVLEAMKMETETINAPDGEVYVYKRDVAVQGGGLIKI 121
 DB 534 WKVIATGOSVAGDVLLILEAMKMETETIRAAQAGTVRGIAVSGDAVSVDITMTL 590

RESULT 20
 AB0909
 oxalacetate decarboxylase (EC 4.1.1.3) alpha chain [imported] - Salmonella enterica
 C:Species: Salmonella enterica subsp. enterica serovar Typhi
 A:Note: this species has also been called Salmonella typhi
 C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 27-Nov-2001
 C:Accession: AB0909
 R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Church

th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, S.; Moule, S.; O'Gaora, P.
 Nature 413, 848-852, 2001
 A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, R.; A:Title: Complete genome sequence of a multiple drug resistant *Salmonella enterica* serovar
 A:Reference number: AB0502; PMID:11677608
 A:Accession: AE0909
 A:Status: Preliminary
 A:Molecule type: DNA
 A:Residues: 1-591 <PAR>
 A:Cross-references: GB:AL513382; PIDN:CAD07867.1; PID:g1504414; GSPDB:GN00176
 C:Genetics:
 A:Gene: oadA
 C:Superfamily: Klebsiella pneumoniae oxaloacetate decarboxylase alpha chain; lipoyl/biotin
 C:Keywords: carbon-carbon lyase; carboxy-lyase

Query Match 28.2%; Score 173.5; DB 2; Length 591;
 Best Local Similarity 40.2%; Pred. No. 2e-06;
 Matches 47; Conservative 14; Mismatches 49; Indels 7; Gaps 3;

QY 5 VYVNGTAVDVVDVDSKSHENPMGTILFGGGTGAPAPAGAGAGAKAGEEIPAPLAGTV 64
 Db 481 VEVGKAFVVRVS-DGGDISQLTFAV--PAASSAPVQAAAPAGAGT---PVTAAPLAGNI 533

QY 65 SKIIVKGGDYVKGQTVLVEAMKMETEINAPTDGKVEKLVKRDVAVGGGGLIKI 121
 Db 534 WKVATGEGVAGEDVLLIEAMKMETEIRAAQAGTVRGIAVKGSDAVSVGDTLMTL 590

RESULT 21

F64453
 oxaloacetate decarboxylase (EC 4.1.1.3) alpha chain M01231 [similarity] - *Methanococcus*
 C:Species: *Methanococcus jannaschii*
 C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 18-Aug-2000

A:Accession: F64453
 R:Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake,
 Reich, C.T.; Overbeek, R.; Kirkness, E.F.; Weissbrock, K.G.; Merrick, J.M.; Glodek, A.;
 rson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.
 Science 273, 1058-1073, 1996

A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C
 A:Title: Complete genome sequence of the methanogenic archaeon, *Methanococcus jannaschii*
 A:Reference number: A64300; M01231; PMID:8688087
 A:Accession: F64453
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-567 <RUL>
 A:Cross-references: GB:U67563; GB:L77117; NID:g2826379; PIDN:AA99233.1; PID:g1591862; E
 C:Genetics:

A:Map position: REV1175190-1173487
 C:Superfamily: Klebsiella pneumoniae oxaloacetate decarboxylase alpha chain; lipoyl/biotin
 C:Keywords: carbon-carbon lyase; carboxy-lyase
 F:494-567/Domain: lipoyl/biotin-binding homology <LPB>
 F:533/Binding site: biotin (lys) (covalent) #status predicted

Query Match 28.1%; Score 173; DB 2; Length 567;
 Best Local Similarity 50.7%; Pred. No. 2.1e-06;
 Matches 35; Conservative 12; Mismatches 22; Indels 0; Gaps 0;

QY 53 EGETPALAGVSKILVKEGDTVAKGQTVLVEAMKMETEINAPTDKVERVLKREDAV 112
 Db 498 EGAVTSFPRGKWTIKIKKEDGKAKGDIYVLEAMKHEPIESVEGIVETILIDEGAV 557

QY 113 QGGGGLIKI 121
 Db 558 NVGDVIMII 566

RESULT 22

D97227
 pyruvate carboxylase, PYKA [imported] - *Clostridium acetobutylicum*
 C:Species: *Clostridium acetobutylicum*
 C:Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 30-Sep-2001
 C:Accession: D97227

R.Nolling, J.; Betton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; L
 ; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
 J. Bacteriol. 183, 4823-4838, 2001
 A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium
 A:Reference number: A96900; M01231; PMID:21359325
 A:Accession: D97227
 A:Status: Preliminary
 A:Molecule type: DNA
 A:Residues: 1-1144 <KUR>
 A:Cross-references: GB:AE001437; PIDN:AAK80607.1; PID:g15025689; GSPDB:GN00168
 A:Experimental source: *Clostridium acetobutylicum* ATCC824
 C:Genetics:
 A:Gene: CAC2660
 C:Superfamily: pyruvate carboxylase; biotin carboxylase homology; lipoyl/biotin-bind

Query Match 27.6%; Score 170; DB 2; Length 1144;
 Best Local Similarity 47.3%; Pred. No. 7.3e-06;
 Matches 35; Conservative 13; Mismatches 26; Indels 0; Gaps 0;

QY 48 AGKAGEEIPAPLAGTVSKILVKEGDTVAKGQTVLVEAMKMETEINAPTDGKVEKLVK 107
 Db 1070 ADSNKKKEIGASIRGNVVFVKRDKVKKGDSLMVTEAMKMETVSVSDGTVGGIFVK 1129

QY 108 ERDAVGGGGLIKI 121
 Db 1130 EGDVQSGGLIKL 1143

RESULT 23

F82966
 Probable transcarboxylase subunit PA5435 [imported] - *Pseudomonas aeruginosa* (strain
 C:Species: *Pseudomonas aeruginosa*
 C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000

A:Accession: F82966
 R:Stover, C.K.; Pham, X.Q.; Ervin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.;
 adman, S.; Yun, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; L
 ; Loay, S.; Olson, M.V.
 Nature 406, 959-964, 2000
 A:Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic pa
 A:Reference number: A82950; M01231; PMID:10984043
 A:Accession: F82966
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-607 <STO>
 A:Cross-references: GB:AE004956; GB:AE004091; NID:g9951760; PIDN:AA08820.1; GSPDB:GN
 A:Experimental source: strain PA01
 C:Genetics:
 A:Gene: PA5435
 C:Superfamily: Klebsiella pneumoniae oxaloacetate decarboxylase alpha chain; lipoyl/b

Query Match 27.4%; Score 168.5; DB 2; Length 607;
 Best Local Similarity 32.4%; Pred. No. 5.1e-06;
 Matches 44; Conservative 20; Mismatches 43; Indels 29; Gaps 3;

QY 5 VYVNGTAVDVVDV-----DVDSKSH-----NPMGTLFGGGTGAPAPAGG 45
 Db 480 IDVHGERTYVDITGVKSDNRKRFYLSIDCMPREVFEPLENVVAGSASGRKHA----- 534

QY 46 AGAKAAGEEIPAPLAGTVSKILVKEGDTVAKGQTVLVEAMKMETEINAPTDKVERVLK 105
 Db 535 -----SEPGHVTSTPMNIDVVLVEKSDVAKGAVLITEAMKMETEIVGAGIAGTVAIH 589

QY 106 VKERDAVGGGGLIKI 121
 Db 590 VAKGDVNPGEILIEI 605

RESULT 24

F70439
 oxaloacetate decarboxylase alpha chain - *Aquifex aeolicus*
 C:Species: *Aquifex aeolicus*
 C:Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 18-Aug-2000
 C:Accession: F70439

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 1, 2003, 07:49:49 ; Search time 173.469 Seconds

(without alignments)
453.439 Million cell updates/sec

Title: US-09-987-485-1

Perfect score: 615
Sequence: 1 MKLKVTVNGTAYDVVDVXK.....KVLVKERDAVQGGGLIKIG 122Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 4569144 seqs, 644733110 residues

Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : Pending_Patents_AA_Main.*

1: /cgn2_6/ptodata/1/paa/PCRTUS_COMB.pep.*
2: /cgn2_6/ptodata/1/paa/US06_COMB.pep.*
3: /cgn2_6/ptodata/1/paa/US07_COMB.pep.*
4: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
5: /cgn2_6/ptodata/1/paa/US081_COMB.pep.*
6: /cgn2_6/ptodata/1/paa/US082_COMB.pep.*
7: /cgn2_6/ptodata/1/paa/US083_COMB.pep.*
8: /cgn2_6/ptodata/1/paa/US084_COMB.pep.*
9: /cgn2_6/ptodata/1/paa/US085_COMB.pep.*
10: /cgn2_6/ptodata/1/paa/US086_COMB.pep.*
11: /cgn2_6/ptodata/1/paa/US087_COMB.pep.*
12: /cgn2_6/ptodata/1/paa/US088_COMB.pep.*
13: /cgn2_6/ptodata/1/paa/US089_COMB.pep.*
14: /cgn2_6/ptodata/1/paa/US090_COMB.pep.*
15: /cgn2_6/ptodata/1/paa/US091_COMB.pep.*
16: /cgn2_6/ptodata/1/paa/US092_COMB.pep.*
17: /cgn2_6/ptodata/1/paa/US093_COMB.pep.*
18: /cgn2_6/ptodata/1/paa/US094_COMB.pep.*
19: /cgn2_6/ptodata/1/paa/US095_COMB.pep.*
20: /cgn2_6/ptodata/1/paa/US096_COMB.pep.*
21: /cgn2_6/ptodata/1/paa/US097_COMB.pep.*
22: /cgn2_6/ptodata/1/paa/US098_COMB.pep.*
23: /cgn2_6/ptodata/1/paa/US099_COMB.pep.*
24: /cgn2_6/ptodata/1/paa/US100_COMB.pep.*
25: /cgn2_6/ptodata/1/paa/US101_COMB.pep.*
26: /cgn2_6/ptodata/1/paa/US102_COMB.pep.*
27: /cgn2_6/ptodata/1/paa/US60_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	615	100.0	122	23	US-09-987-485-1
2	615	100.0	126	15	US-09-148-920-5
3	615	100.0	129	21	US-09-791-537-120787
4	615	100.0	133	26	US-10-251-313-7
5	615	100.0	256	15	US-09-148-920-3
6	615	100.0	266	15	US-09-148-920-7

7	604.5	98.3	123	21	US-09-791-537-132235	Sequence 132235,
8	604.5	98.3	123	26	US-10-251-313-1	Sequence 1, Appl1
9	377	61.3	77	21	US-09-791-537-49119	Sequence 49119, A
10	342	55.6	70	23	US-09-987-485-2	Sequence 2, Appl1
11	231.5	37.6	145	21	US-09-791-537-111788	Sequence 111788,
12	221	35.9	149	21	US-09-791-537-111800	Sequence 111800,
13	205	33.3	129	21	US-09-791-537-54812	Sequence 54812, A
14	204.5	33.3	589	27	US-09-791-537-10079	Sequence 10079, A
15	203.5	33.6	146	19	US-09-540-2098-9732	Sequence 9732, Ap
16	200.5	32.6	596	21	US-09-791-537-72996	Sequence 72996, A
17	198.5	32.3	145	21	US-09-791-537-5404	Sequence 5404, Ap
18	186.5	32.0	145	21	US-09-791-537-48753	Sequence 48753, A
19	180.5	31.0	655	21	US-09-791-537-31887	Sequence 31887, A
20	190.5	31.0	645	27	US-09-791-537-111780	Sequence 111780,
21	190	30.9	150	21	US-09-360-039-112	Sequence 112, App
22	187.5	30.5	599	21	US-09-791-537-36890	Sequence 36890, A
23	184	29.9	571	21	US-09-791-537-31867	Sequence 31867, A
24	184	29.9	571	27	US-09-360-039-1263	Sequence 1263, Ap
25	184	29.9	571	27	US-09-360-039-20329	Sequence 20329, A
26	180.5	29.3	1076	27	US-09-360-039-13831	Sequence 13831, A
27	176.5	28.7	574	27	US-09-360-039-14349	Sequence 14349, A
28	176.5	28.7	574	27	US-09-360-039-15017	Sequence 15017, A
29	176.5	28.7	575	27	US-09-360-039-11591	Sequence 11591, A
30	176.5	28.7	575	27	US-09-360-039-14608	Sequence 14608, A
31	176.5	28.7	576	22	US-09-739-449-11573	Sequence 11573, A
32	176.5	28.7	576	22	US-09-803-110-11573	Sequence 11573, A
33	176	28.6	134	21	US-09-791-537-111798	Sequence 111798,
34	176	28.6	596	21	US-09-791-537-107230	Sequence 107230,
35	176	28.6	984	21	US-09-791-537-24359	Sequence 24359, A
36	175.5	28.5	147	15	US-09-134-000-6221	Sequence 6221, Ap
37	175.5	28.5	591	21	US-09-791-537-121743	Sequence 121743,
38	175.5	28.5	610	27	US-09-360-039-18634	Sequence 18634, A
39	173	28.1	567	21	US-09-791-537-46258	Sequence 46258, A
40	173	28.1	567	27	US-09-360-039-1053	Sequence 1053, Ap
41	170.5	27.7	591	20	US-09-602-787A-24	Sequence 24, Appl
42	170.5	27.7	591	21	US-09-738-626-6940	Sequence 6940, Ap
43	170.5	27.7	591	21	US-09-791-537-26361	Sequence 26361, A
44	170	27.6	169	23	US-09-902-540-16044	Sequence 16044, A
45	168.5	27.4	612	16	US-09-252-991A-19134	Sequence 19134, A
46	168.5	27.4	620	27	US-09-791-537-111784	Sequence 111784,
47	168.5	27.4	620	27	US-09-360-039-119	Sequence 119, App
48	168	27.3	130	21	US-09-791-537-132236	Sequence 132236,
49	166	27.0	600	18	US-09-415-884-26	Sequence 26, Appl
50	166	27.0	600	21	US-09-791-537-1248	Sequence 1248, Ap
51	165	26.8	1150	27	US-09-360-039-17338	Sequence 17338, A
52	164	26.7	654	27	US-09-360-039-16934	Sequence 16934, A
53	163	26.5	142	21	US-09-791-537-111784	Sequence 111784,
54	163	26.5	598	21	US-09-791-537-7608	Sequence 7608, Ap
55	163	26.5	598	21	US-09-791-537-64192	Sequence 64192, A
56	162.5	26.4	1157	23	US-09-974-973-2	Sequence 2, Appl1
57	162.5	26.4	1157	23	US-09-974-973-4	Sequence 10553,
58	162	26.3	591	21	US-09-791-537-105564	Sequence 105564,
59	162	26.3	599	21	US-09-791-537-105563	Sequence 105563,
60	161.5	26.3	1092	21	US-09-739-449-11312	Sequence 11312, A
61	161.5	26.3	1092	22	US-09-803-110-11312	Sequence 11312, A
62	161.5	26.3	1147	27	US-09-360-039-11450	Sequence 11450, A
63	161.5	26.3	1151	27	US-09-360-039-14817	Sequence 14817, A
64	161.5	26.3	1151	27	US-09-360-039-15006	Sequence 15006, A
65	159.5	25.9	272	20	US-09-602-740-112	Sequence 112, App
66	159.5	25.9	272	20	US-09-602-740-114	Sequence 114, App
67	159.5	25.9	1140	19	US-09-529-043A-2	Sequence 2, Appl1
68	159.5	25.9	1140	19	US-09-529-043B-2	Sequence 2, Appl1
69	159.5	25.9	1140	21	US-09-738-626-4265	Sequence 4265, Ap
70	159.5	25.9	1140	21	US-09-791-537-45746	Sequence 45746, A
71	159.5	25.9	1140	23	US-09-974-973-19	Sequence 19, Appl
72	159.5	25.9	1140	24	US-10-045-072-2	Sequence 2, Appl1
73	159	25.9	161	19	US-09-583-110-3907	Sequence 3907, Ap
74	159	25.9	161	21	US-09-752-069A-198	Sequence 198, App
75	159	25.9	161	27	US-09-174-089-198	Sequence 198, App
76	159	25.9	163	1	PCT-US97-14436-402	Sequence 402, App
77	159	25.9	163	13	US-08-911-503A-402	Sequence 402, App
78	159	25.9	163	13	US-08-911-503A-402	Sequence 402, App
79	159	25.9	163	13	US-09-107-433-3683	Sequence 3683, Ap

80 158.5 25.8 593 21 US-09-791-537-31883 Sequence 31883, A
81 157 25.5 161 1 PCT-US02-03987-13363 Sequence 13363, A
82 157 25.5 161 1 PCT-US02-03987-13363 Sequence 13363, A
83 157 25.5 161 22 US-09-815-242-13363 Sequence 13363, A
84 157 25.5 161 22 US-09-815-242-13363 Sequence 13363, A
85 157 25.5 161 24 US-10-072-851-13363 Sequence 13363, A
86 157 25.5 161 24 US-10-072-851-13363 Sequence 13363, A
87 156 25.4 989 12 US-08-827-356-5175 Sequence 5175, A
88 156 25.4 989 20 US-09-611-529-3973 Sequence 3973, A
89 156 25.4 1073 1 PCT-US02-03987-12361 Sequence 12361, A
90 156 25.4 1073 22 US-09-815-242-12361 Sequence 12361, A
91 156 25.4 1073 22 US-10-072-851-12361 Sequence 12361, A
92 156 25.4 1147 2 PCT-US02-03987-5468 Sequence 5468, A
93 156 25.4 1147 22 US-09-815-242-5468 Sequence 5468, A
94 156 25.4 1147 24 US-10-072-851-5468 Sequence 5468, A
95 156 25.4 1147 27 US-60-253-625-2492 Sequence 2492, A
96 156 25.4 1147 27 US-60-257-931-3012 Sequence 3012, A
97 156 25.4 1147 27 US-60-269-308-4033 Sequence 4033, A
98 156 25.2 1156 25 US-10-151-129-23 Sequence 23, A
99 155 25.2 167 21 US-09-791-537-63783 Sequence 63783, A
100 154.5 25.1 1124 21 US-09-791-537-66475 Sequence 66475, A

ALIGNMENTS

RESULT 1
US-09-987-485-1
; Sequence 1, Application US/09987485
; GENERAL INFORMATION:
; APPLICANT: Barry, Michael
; TITLE OF INVENTION: Methods for the In Vivo Biotin Labeling of Polypeptides
; FILE REFERENCE: 15987/268653
; CURRENT APPLICATION NUMBER: US/09/987,485
; CURRENT FILING DATE: 2001-11-14
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 122
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-987-485-1

Query Match 100.0%; Score 615; DB 23; Length 122;
Best Local Similarity 100.0%; Pred. No. 7.7e-52;
Matches 122; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKLKVYNGTAYVDVVDKSHENPMGTILFGGCTGAPAPAGAGAGAGGEGEIPAPL 60
DB 1 MKLKVYNGTAYVDVVDKSHENPMGTILFGGCTGAPAPAGAGAGAGGEGEIPAPL 60
QY 61 AGTYSKILVEGDTYKAGOTVLYLEAMKMETEINAPTDGKVEKVLKERDAVGGGGLIK 120
DB 61 AGTYSKILVEGDTYKAGOTVLYLEAMKMETEINAPTDGKVEKVLKERDAVGGGGLIK 120
QY 121 IG 122
DB 121 IG 122

RESULT 2
US-09-148-920-5
; Sequence 5, Application US/09148920
; GENERAL INFORMATION:
; APPLICANT: Mullerix, Michael C.
; APPLICANT: Deusch, John
; TITLE OF INVENTION: Recombinant Antigen Immunoassay for the Diagnosis of
; FILE REFERENCE: P-4131
; CURRENT APPLICATION NUMBER: US/09/148,920
; CURRENT FILING DATE: 1997-09-04
; NUMBER OF SEQ ID NOS: 7

SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 126
; TYPE: PRT
; ORGANISM: Treponema pallidum
US-09-148-920-5

Query Match 100.0%; Score 615; DB 15; Length 126;
Best Local Similarity 100.0%; Pred. No. 8e-52;
Matches 122; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKLKVYNGTAYVDVVDKSHENPMGTILFGGCTGAPAPAGAGAGAGGEGEIPAPL 60
DB 1 MKLKVYNGTAYVDVVDKSHENPMGTILFGGCTGAPAPAGAGAGAGGEGEIPAPL 60
QY 61 AGTYSKILVEGDTYKAGOTVLYLEAMKMETEINAPTDGKVEKVLKERDAVGGGGLIK 120
DB 61 AGTYSKILVEGDTYKAGOTVLYLEAMKMETEINAPTDGKVEKVLKERDAVGGGGLIK 120
QY 121 IG 122
DB 121 IG 122

RESULT 3
US-09-791-537-120787
; Sequence 120787, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Biomolix, Inc.
; APPLICANT: Debe, Derek
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 120787
; LENGTH: 129
; TYPE: PRT
; ORGANISM: Cloning vector pInPoint
US-09-791-537-120787

Query Match 100.0%; Score 615; DB 21; Length 129;
Best Local Similarity 100.0%; Pred. No. 8.3e-52;
Matches 122; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKLKVYNGTAYVDVVDKSHENPMGTILFGGCTGAPAPAGAGAGAGGEGEIPAPL 60
DB 1 MKLKVYNGTAYVDVVDKSHENPMGTILFGGCTGAPAPAGAGAGAGGEGEIPAPL 60
QY 61 AGTYSKILVEGDTYKAGOTVLYLEAMKMETEINAPTDGKVEKVLKERDAVGGGGLIK 120
DB 61 AGTYSKILVEGDTYKAGOTVLYLEAMKMETEINAPTDGKVEKVLKERDAVGGGGLIK 120
QY 121 IG 122
DB 121 IG 122

RESULT 4
US-10-251-313-7
; Sequence 7, Application US/10251313
; GENERAL INFORMATION:
; APPLICANT: AMBROSIS, DOROTHEE
; APPLICANT: LANZENDORFER, MARTIN
; APPLICANT: SCHRAEM, MICHAEL
; APPLICANT: WATZEL, MANFRED
; TITLE OF INVENTION: IMPROVED METHOD FOR A SEQUENCE SPECIFIC BIOTINYLATION
; FILE REFERENCE: 506.1001
; CURRENT APPLICATION NUMBER: US/10/251,313
; CURRENT FILING DATE: 2002-09-20

PRIOR APPLICATION NUMBER: EP 01123681.1
PRIOR FILING DATE: 2001-12-13
PRIOR APPLICATION NUMBER: EP 01122554.7
PRIOR FILING DATE: 2001-09-25
NUMBER OF SEQ ID NOS: 14
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 7
LENGTH: 133
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Pinpoint
OTHER INFORMATION: Biotinylation peptide
US-10-251-313-7

Query Match 100.0%; Score 615; DB 26; Length 133;
Best Local Similarity 100.0%; Pred. No. 8.6e-52;
Matches 122; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKLKYVNGTAYDVVDVVDKSHENPMGTILFGGCGAPAPAGAGAGAGGEGEIPAPL 60
DB 1 MKLKYVNGTAYDVVDVVDKSHENPMGTILFGGCGAPAPAGAGAGAGGEGEIPAPL 60
QY 61 AGTYSKILVREGDVKAGQTVLVEAMKMETEINAPTDGKVEKYLKERDAVGGGGLIK 120
DB 61 AGTYSKILVREGDVKAGQTVLVEAMKMETEINAPTDGKVEKYLKERDAVGGGGLIK 120
QY 121 IG 122
DB 121 IG 122

RESULT 5
US-09-148-920-3
Sequence 3, Application US/09148920
GENERAL INFORMATION:
APPLICANT: Mullenix, Michael C.
TITLE OF INVENTION: Recombinant Antigen Immunoassay for the Diagnosis of
TITLE OF INVENTION: Syphilis
FILE REFERENCE: P-4131
CURRENT APPLICATION NUMBER: US/09/148,920
CURRENT FILING DATE: 1997-09-04
NUMBER OF SEQ ID NOS: 7
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 3
LENGTH: 256
TYPE: PRT
ORGANISM: Treponema pallidum
US-09-148-920-3

Query Match 100.0%; Score 615; DB 15; Length 256;
Best Local Similarity 100.0%; Pred. No. 2e-51;
Matches 122; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKLKYVNGTAYDVVDVVDKSHENPMGTILFGGCGAPAPAGAGAGAGGEGEIPAPL 60
DB 1 MKLKYVNGTAYDVVDVVDKSHENPMGTILFGGCGAPAPAGAGAGAGGEGEIPAPL 60
QY 61 AGTYSKILVREGDVKAGQTVLVEAMKMETEINAPTDGKVEKYLKERDAVGGGGLIK 120
DB 61 AGTYSKILVREGDVKAGQTVLVEAMKMETEINAPTDGKVEKYLKERDAVGGGGLIK 120
QY 121 IG 122
DB 121 IG 122

RESULT 6
US-09-148-920-7
Sequence 7, Application US/09148920
GENERAL INFORMATION:
APPLICANT: Mullenix, Michael C.

APPLICANT: Deutsch, John
TITLE OF INVENTION: Recombinant Antigen Immunoassay for the Diagnosis of
TITLE OF INVENTION: Syphilis
FILE REFERENCE: P-4131
CURRENT APPLICATION NUMBER: US/09/148,920
CURRENT FILING DATE: 1997-09-04
NUMBER OF SEQ ID NOS: 7
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 7
LENGTH: 266
TYPE: PRT
ORGANISM: Treponema pallidum
US-09-148-920-7

Query Match 100.0%; Score 615; DB 15; Length 266;
Best Local Similarity 100.0%; Pred. No. 2.1e-51;
Matches 122; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKLKYVNGTAYDVVDVVDKSHENPMGTILFGGCGAPAPAGAGAGAGGEGEIPAPL 60
DB 1 MKLKYVNGTAYDVVDVVDKSHENPMGTILFGGCGAPAPAGAGAGAGGEGEIPAPL 60
QY 61 AGTYSKILVREGDVKAGQTVLVEAMKMETEINAPTDGKVEKYLKERDAVGGGGLIK 120
DB 61 AGTYSKILVREGDVKAGQTVLVEAMKMETEINAPTDGKVEKYLKERDAVGGGGLIK 120
QY 121 IG 122
DB 121 IG 122

RESULT 7
US-09-791-537-132235
Sequence 132235, Application US/09791537
GENERAL INFORMATION:
APPLICANT: Biomolix, Inc.
APPLICANT: Debe, Derek
TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
TITLE OF INVENTION: METHODS OF USE THEREOF
FILE REFERENCE: 261/210
CURRENT APPLICATION NUMBER: US/09/791,537
CURRENT FILING DATE: 2001-02-22
NUMBER OF SEQ ID NOS: 153055
SOFTWARE: Patentin version 3.0
SEQ ID NO 132235
LENGTH: 123
TYPE: PRT
ORGANISM: Propionibacterium freudenreichii subsp
US-09-791-537-132235

Query Match 98.3%; Score 604.5; DB 21; Length 123;
Best Local Similarity 99.2%; Pred. No. 8.2e-51;
Matches 122; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 MKLKYVNGTAYDVVDVVDKSHENPMGTILFGGCGAPAPAGAGAGAGGEGEIPAPL 59
DB 1 MKLKYVNGTAYDVVDVVDKSHENPMGTILFGGCGAPAPAGAGAGAGGEGEIPAPL 60
QY 60 LAGTYSKILVREGDVKAGQTVLVEAMKMETEINAPTDGKVEKYLKERDAVGGGGLI 119
DB 61 LAGTYSKILVREGDVKAGQTVLVEAMKMETEINAPTDGKVEKYLKERDAVGGGGLI 120
QY 120 KIG 122
DB 121 KIG 123

RESULT 8
US-10-251-313-1
Sequence 1, Application US/10251313
GENERAL INFORMATION:
APPLICANT: AMBROSIOUS, DOROTHEE

APPLICANT: LANZENDORFER, MARTIN
APPLICANT: SCHRAEML, MICHAEL
APPLICANT: MATZELE, MANFRED
TITLE OF INVENTION: IMPROVED METHOD FOR A SEQUENCE SPECIFIC BIOTINYLATION
FILE REFERENCE: 506.1001
CURRENT APPLICATION NUMBER: US/10/251.313
CURRENT FILING DATE: 2002-09-20
PRIOR APPLICATION NUMBER: EP 01129681.1
PRIOR FILING DATE: 2001-12-13
PRIOR APPLICATION NUMBER: EP 01122554.7
PRIOR FILING DATE: 2001-09-25
NUMBER OF SEQ ID NOS: 14
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 123
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: 1.35
OTHER INFORMATION: transcarboxylase subunit of Propionibacterium
OTHER INFORMATION: Sherman1
US-10-251-313-1

Query Match 98.3%; Score 604.5; DB 26; Length 123;
Best Local Similarity 99.2%; Pred. No. 8.2e-51;
Matches 122; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

OY 1 MKLKVTVNGTAVDNDVDKSHENPMGTILFGGTGAPAP-AAGAGAGAGGEGEIPAP 59
DB 1 MKLKVTVNGTAVDNDVDKSHENPMGTILFGGTGAPAPRAAGAGAGAGGEGEIPAP 60
OY 60 LAGTVSKILYKEDGYKAGQTVLVLEAMKMETEINAPTDGKVEKVLKEDAVOGGGLI 119
DB 61 LAGTVSKILYKEDGYKAGQTVLVLEAMKMETEINAPTDGKVEKVLKEDAVOGGGLI 120
OY 120 KIG 122
DB 121 KIG 123

RESULT 9
US-09-791-537-49119
Sequence 49119, Application US/09791537
GENERAL INFORMATION:
APPLICANT: Bionomix, Inc.
APPLICANT: Debe, Derek
APPLICANT: Danzer, Joseph
TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
FILE REFERENCE: 261/210
CURRENT APPLICATION NUMBER: US/09/791.537
CURRENT FILING DATE: 2001-02-22
NUMBER OF SEQ ID NOS: 153055
SOFTWARE: PatentIn version 3.0
SEQ ID NO 49119
LENGTH: 77
TYPE: PRT
ORGANISM: pdb 1DD2A
US-09-791-537-49119

Query Match 61.3%; Score 377; DB 21; Length 77;
Best Local Similarity 100.0%; Pred. No. 7.4e-29;
Matches 77; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 46 AGAGKAGEGEPAPLAGTVSKILYKEDGYKAGQTVLVLEAMKMETEINAPTDGKVEKVL 105
DB 1 AGAGKAGEGEPAPLAGTVSKILYKEDGYKAGQTVLVLEAMKMETEINAPTDGKVEKVL 60
OY 106 VKERDAVVGOGGLIKIG 122
DB 61 VKERDAVVGOGGLIKIG 77

RESULT 10
US-09-987-485-2
Sequence 2, Application US/09987485
GENERAL INFORMATION:
APPLICANT: Barry, Michael
APPLICANT: Parrott, Michael
TITLE OF INVENTION: Methods for the In Vivo Biotin Labeling of Polypeptides
FILE REFERENCE: 15987/268653
CURRENT APPLICATION NUMBER: US/09/987.485
CURRENT FILING DATE: 2001-11-14
NUMBER OF SEQ ID NOS: 5
SOFTWARE: PatentIn version 3.0
SEQ ID NO 2
LENGTH: 70
TYPE: PRT
ORGANISM: Escherichia coli
US-09-987-485-2

Query Match 55.6%; Score 342; DB 23; Length 70;
Best Local Similarity 100.0%; Pred. No. 1.7e-25;
Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 53 EGEIPAPLAGTVSKILYKEDGYKAGQTVLVLEAMKMETEINAPTDGKVEKVLKEDAV 112
DB 1 EGEIPAPLAGTVSKILYKEDGYKAGQTVLVLEAMKMETEINAPTDGKVEKVLKEDAV 60
OY 113 OGOGGLIKIG 122
DB 61 OGOGGLIKIG 70

RESULT 11
US-09-791-537-111788
Sequence 111788, Application US/09791537
GENERAL INFORMATION:
APPLICANT: Bionomix, Inc.
APPLICANT: Debe, Derek
APPLICANT: Danzer, Joseph
TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
FILE REFERENCE: 261/210
CURRENT APPLICATION NUMBER: US/09/791.537
CURRENT FILING DATE: 2001-02-22
NUMBER OF SEQ ID NOS: 153055
SOFTWARE: PatentIn version 3.0
SEQ ID NO 111788
LENGTH: 145
TYPE: PRT
ORGANISM: Pyrococcus abyssi
US-09-791-537-111788

Query Match 37.6%; Score 231.5; DB 21; Length 145;
Best Local Similarity 41.4%; Pred. No. 2.8e-14;
Matches 60; Conservative 18; Mismatches 44; Indels 23; Gaps 3;

OY 1 MKLKVTVNGTAVDNDVDKSHENPMGTILFGGTGAPAP 40
DB 1 MKKVTVNGTAVDNDVDKSHENPMGTILFGGTGAPAP 60
OY 41 PA---AGAGAGAGGEGEIPAPLAGTVSKILYKEDGYKAGQTVLVLEAMKMETEINAP 97
DB 61 PTPVQAPPTPVQVASENVVYAPAPGKVLKILVOGGVXKGLLLEAMKMETEINAP 120
OY 98 DGKVEKVLKEDAVOGGGLIKIG 122
DB 121 DGKVEKVLKEDAVOGGGLIKIG 145

RESULT 12
US-09-791-537-111800
Sequence 111800, Application US/09791537
GENERAL INFORMATION:
APPLICANT: Bionomix, Inc.


```
FILE REFERENCE: 38-10(52052)A
CURRENT APPLICATION NUMBER: US/60/360.039
CURRENT FILING DATE: 2002-02-21
NUMBER OF SEQ ID NOS: 47374
SEQ ID NO 112
LENGTH: 655
TYPE: PRT
ORGANISM: Aquifex aeolicus
US-60-360-039-112

Query Match
Best Local Similarity 31.0%; Score 190.5; DB 27; Length 655;
Matches 49; Conservative 19; Mismatches 52; Indels 13; Gaps 3;

OY 2 KLTAVNGTAYDVVD-----VDVDRKSHENPMGTLFGGCTGAGAPAPAGAGAG 49
DB 513 KFKVIEGVSVDAESGKPRKYVRINNRLKEIQLKPFKEATVQSGAGQVQSAEEEGIP 572
OY 50 RAGE-GEIPAPLAGTAVSKILVEGDTVAKGQTVLLEAMKMETEINAPTDGKVEKVLK 108
DB 573 KATEBGDTVPMPKGVKILVKEGEPVQGGTVATVLEAMKMEVHAPIDGIVKKIFAP 632
OY 109 RPAVGGGGLIKI 121
DB 633 GDVNPDAQIMRI 645

RESULT 21
US-09-791-537-111780
Sequence 111780, Application US/09791537
GENERAL INFORMATION:
APPLICANT: Biomomix, Inc.
APPLICANT: Danzer, Joseph
APPLICANT: Debe, Derek
TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
FILE REFERENCE: 261/210
CURRENT APPLICATION NUMBER: US/09/791.537
CURRENT FILING DATE: 2001-02-22
NUMBER OF SEQ ID NOS: 153055
SOFTWARE: PatentIn version 3.0
SEQ ID NO 111780
LENGTH: 140
TYPE: PRT
ORGANISM: Archaeoglobus fulgidus
US-09-791-537-111780

Query Match
Best Local Similarity 30.9%; Score 190; DB 21; Length 140;
Matches 52; Conservative 19; Mismatches 37; Indels 40; Gaps 4;

OY 1 MKLTAVNGTAYDVVDVDRKSHENPMGTLFGGCTGAGAPAPAGAGAGAGAGAGAG 40
DB 1 MKTVEKVGKKEVEVE-----EVSFK---VEEVKNGKKAIVLEKKFEFEKADIRE 53
OY 41 -----PAAGAGAGAGAGEIPAPLAGTAVSKILVEGDTVAKGQTVLLEAMKMET 91
DB 54 RFAERREBARAEVKATGAKA-----ITAPAGVVTIKLVKGVKVAAGETVLLIEMKME 109
OY 92 ETNAPTDGKVEKVLKVEKPAVGGGLI 119
DB 110 PLASPEDGEIAIVVKEGDKVASGDVLY 137

RESULT 22
US-09-791-537-36890
Sequence 36890, Application US/09791537
GENERAL INFORMATION:
APPLICANT: Biomomix, Inc.
APPLICANT: Danzer, Joseph
APPLICANT: Debe, Derek
TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
FILE REFERENCE: 38-10(52052)A
```

```
FILE REFERENCE: 261/210
CURRENT APPLICATION NUMBER: US/09/791.537
CURRENT FILING DATE: 2001-02-22
NUMBER OF SEQ ID NOS: 153055
SOFTWARE: PatentIn version 3.0
SEQ ID NO 36890
LENGTH: 599
TYPE: PRT
ORGANISM: Campylobacter jejuni
US-09-791-537-36890

Query Match
Best Local Similarity 30.5%; Score 187.5; DB 21; Length 599;
Matches 48; Conservative 15; Mismatches 40; Indels 25; Gaps 3;

OY 2 KLTAVNGTAYDVVD-----VDVDRKSHENPMGTLFGGCTGAGAPAPAGAGAGAG 50
DB 481 KFTVAVNGNKHYHEVSFYFDKDVNKSVMKVEENK--NISSNSTSYDA----- 528
OY 51 AGEGEIPAPLAGTAVSKILVEGDTVAKGQTVLLEAMKMETEINAPTDGKVEKVLK 110
DB 529 --ENEVLAGISGNVPRKIYVNEGEYKSGQALNVLEAMKMEIVNAPKRCIILELCIKIGD 586
OY 111 AVGGGGL 118
DB 587 TVNEGEVL 594

RESULT 23
US-09-791-537-31867
Sequence 31867, Application US/09791537
GENERAL INFORMATION:
APPLICANT: Biomomix, Inc.
APPLICANT: Danzer, Joseph
APPLICANT: Debe, Derek
TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
FILE REFERENCE: 261/210
CURRENT APPLICATION NUMBER: US/09/791.537
CURRENT FILING DATE: 2001-02-22
NUMBER OF SEQ ID NOS: 153055
SOFTWARE: PatentIn version 3.0
SEQ ID NO 31867
LENGTH: 571
TYPE: PRT
ORGANISM: Pyrococcus horikoshii
US-09-791-537-31867

Query Match
Best Local Similarity 29.9%; Score 184; DB 21; Length 571;
Matches 45; Conservative 24; Mismatches 43; Indels 8; Gaps 3;

OY 3 LKVTAVNGTAYDVVDVDRKSHENPMGTLFGGCTGAGAPAPAGAGAGAGAGAGAGAG 61
DB 458 IKIYINGKEFEVEFGIEEPKPRPV---QALPSQPKREVAPSGSV---VSAPWP 510
OY 62 GTVSKILVEGDTVAKGQTVLLEAMKMETEINAPTDGKVEKVLKVEKPAVGGGLIKI 121
DB 511 GKVLAVLVAVGSRVAVGGGLVLEAMKMEIIPSRDGVYKILVKEGSAVDGTGPIIEL 570

RESULT 24
US-60-360-039-1263
Sequence 1263, Application US/60360039
GENERAL INFORMATION:
APPLICANT: Cao, Yongwei
APPLICANT: Chen, Xianfeng
APPLICANT: Goldman, Barry S.
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
FILE REFERENCE: 38-10(52052)A
```

;; CURRENT APPLICATION NUMBER: US/60/360, 039
;; CURRENT FILING DATE: 2002-02-21
;; NUMBER OF SEQ ID NOS: 47374
;; SEQ ID NO 1263
;; LENGTH: 571
;; TYPE: PRT
;; ORGANISM: Pyrococcus horikoshii
US-60-360-039-1263

Query Match 29.9%; Score 184; DB 27; Length 571;
Best Local Similarity 37.5%; Pred. No. 7.3e-09;
Matches 45; Conservative 24; Mismatches 43; Indels 8; Gaps 3;

QY 3 LKTVNGTAVDVVD-VDKSHENPMGTILFGGTGAPAPAGAGAGKAGEEIPAPLA 61
DB 453 IKIYNGKEFEVEEGIEFEPKPKPV---QAIPOPKREVVAAPSGSV---VSAPMP 510
QY 62 GTVSKILVKEGDTYKAGQTVLVLEAMKETEINAPTDGKVEKVLVKERDAVGGGGLIKI 121
DB 511 GKVLRLVLRVGDVRVGGGLVLEAMKEMENIPSPROGVYRIILVKEGAVDTGQPLIEL 570

RESULT 25
US-60-360-039-20329
;; Sequence 20329, Application US/60360039
;; GENERAL INFORMATION:

;; APPLICANT: Cao, Yongwei
;; APPLICANT: Chen, Xianfeng
;; APPLICANT: Goldman, Barry S.
;; APPLICANT: Hinkle, Gregory J.
;; APPLICANT: Slater, Steven C.
;; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
;; FILE REFERENCE: 38-10(52052)A
;; CURRENT APPLICATION NUMBER: US/60/360, 039
;; CURRENT FILING DATE: 2002-02-21
;; NUMBER OF SEQ ID NOS: 47374
;; SEQ ID NO 20329
;; LENGTH: 571
;; TYPE: PRT
;; ORGANISM: Pyrococcus horikoshii
US-60-360-039-20329

Query Match 29.9%; Score 184; DB 27; Length 571;
Best Local Similarity 37.5%; Pred. No. 7.3e-09;
Matches 45; Conservative 24; Mismatches 43; Indels 8; Gaps 3;

QY 3 LKTVNGTAVDVVD-VDKSHENPMGTILFGGTGAPAPAGAGAGKAGEEIPAPLA 61
DB 458 IKIYNGKEFEVEEGIEFEPKPKPV---QAIPOPKREVVAAPSGSV---VSAPMP 510
QY 62 GTVSKILVKEGDTYKAGQTVLVLEAMKETEINAPTDGKVEKVLVKERDAVGGGGLIKI 121
DB 511 GKVLRLVLRVGDVRVGGGLVLEAMKEMENIPSPROGVYRIILVKEGAVDTGQPLIEL 570

RESULT 26
US-60-360-039-13831
;; Sequence 13831, Application US/60360039
;; GENERAL INFORMATION:

;; APPLICANT: Cao, Yongwei
;; APPLICANT: Chen, Xianfeng
;; APPLICANT: Goldman, Barry S.
;; APPLICANT: Hinkle, Gregory J.
;; APPLICANT: Slater, Steven C.
;; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
;; FILE REFERENCE: 38-10(52052)A
;; CURRENT APPLICATION NUMBER: US/60/360, 039
;; CURRENT FILING DATE: 2002-02-21
;; NUMBER OF SEQ ID NOS: 47374
;; SEQ ID NO 13831
;; LENGTH: 1076

;; TYPE: PRT
;; ORGANISM: Pseudomonas fluorescens
US-60-360-039-13831

Query Match 29.3%; Score 180.5; DB 27; Length 1076;
Best Local Similarity 34.6%; Pred. No. 3.7e-08;
Matches 47; Conservative 17; Mismatches 43; Indels 29; Gaps 3;

QY 5 VTNGTAVDVVDV-----DVDKSHE-----NPMGTILFGGTGAPAPAGC 45
DB 950 IDVGHETRYVDITGVYKAEGRHFYLSIDOMPEVVEFEPLENEFVGSSSRKQASA--- 1006
QY 46 AGAKAGEEIPAPLACTVSKILVKEGDTYKAGQTVLVLEAMKETEINAPTDGKVEKVL 105
DB 1007 -----PGHVSITPMGNIDVLVKEGDTYKAGQAVLITAMKETEVOAIAIGKVTAIH 1059
QY 106 VKERDAVGGGGLIKI 121
DB 1060 VAKGDRVNPGEILIEL 1075

RESULT 27
US-60-360-039-14349
;; Sequence 14349, Application US/60360039
;; GENERAL INFORMATION:

;; APPLICANT: Cao, Yongwei
;; APPLICANT: Chen, Xianfeng
;; APPLICANT: Goldman, Barry S.
;; APPLICANT: Hinkle, Gregory J.
;; APPLICANT: Slater, Steven C.
;; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
;; FILE REFERENCE: 38-10(52052)A
;; CURRENT APPLICATION NUMBER: US/60/360, 039
;; CURRENT FILING DATE: 2002-02-21
;; NUMBER OF SEQ ID NOS: 47374
;; SEQ ID NO 14349
;; LENGTH: 574
;; TYPE: PRT
;; ORGANISM: Agrobacterium tumefaciens
US-60-360-039-14349

Query Match 28.7%; Score 176.5; DB 27; Length 574;
Best Local Similarity 43.8%; Pred. No. 4e-08;
Matches 42; Conservative 17; Mismatches 28; Indels 9; Gaps 3;

QY 26 MGTILFGGTGAPAPAPAGAGAGKAGEEIPAPLAGTYSKILVKEGDTYKAGQTVLVLE 85
DB 488 LGTV--SGNAPSANV-----EKREGMTAPVSGTDSFKYKDGETVSEGDILAVME 539
QY 86 AMKETEINAPTDGKVEKVLVKERDAVGGGGLIKI 121
DB 540 AMKETEIVATVRAGKV-RLIVKEGDTYKAGQAVLIDI 574

RESULT 28
US-60-360-039-15017
;; Sequence 15017, Application US/60360039
;; GENERAL INFORMATION:

;; APPLICANT: Cao, Yongwei
;; APPLICANT: Chen, Xianfeng
;; APPLICANT: Goldman, Barry S.
;; APPLICANT: Hinkle, Gregory J.
;; APPLICANT: Slater, Steven C.
;; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
;; FILE REFERENCE: 38-10(52052)A
;; CURRENT APPLICATION NUMBER: US/60/360, 039
;; CURRENT FILING DATE: 2002-02-21
;; NUMBER OF SEQ ID NOS: 47374
;; SEQ ID NO 15017
;; LENGTH: 574
;; TYPE: PRT

ORGANISM: Agrobacterium tumefaciens
US-60-360-039-15017

Query Match 28.7%; Score 176.5; DB 27; Length 574;
Best Local Similarity 43.8%; Pred. No. 4e-08;
Matches 42; Conservative 17; Mismatches 28; Indels 9; Gaps 3;

OY 26 MGTILFGGTTGAPAPAGAGAGAGGEGTIPAPLAGTIVSKILYKEDGTVKAGOTVLYLE 85
DB 488 LGTV--SGGNASAPSAV-----EKKEGEMTAPVSGTLQSFVKDGETVSEGDLLAYME 539

OY 86 AMKMETEINAPTDGKVEKVLKERDAVVOGGGLIKI 121
DB 540 AMKMETQIVATRAKGV-RLIYKEDGYLQAGATLIDI 574

RESULT 29
US-60-360-039-11591
Sequence 11591, Application US/60360039

GENERAL INFORMATION:

APPLICANT: Cao, Yongwei

APPLICANT: Goldman, Barry S.

APPLICANT: Hinkle, Gregory J.

APPLICANT: Slater, Steven C.

TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF

FILE REFERENCE: 38-10(52052)A

CURRENT APPLICATION NUMBER: US/60/360,039

CURRENT FILING DATE: 2002-02-21

NUMBER OF SEQ ID NOS: 47374

SEQ ID NO 11591

LENGTH: 575

TYPE: PRT

ORGANISM: Agrobacterium tumefaciens

US-60-360-039-11591

Query Match 28.7%; Score 176.5; DB 27; Length 575;
Best Local Similarity 43.8%; Pred. No. 4e-08;
Matches 42; Conservative 17; Mismatches 28; Indels 9; Gaps 3;

OY 26 MGTILFGGTTGAPAPAGAGAGAGGEGTIPAPLAGTIVSKILYKEDGTVKAGOTVLYLE 85
DB 489 LGTV--SGGNASAPSAV-----EKKEGEMTAPVSGTLQSFVKDGETVSEGDLLAYME 540

OY 86 AMKMETEINAPTDGKVEKVLKERDAVVOGGGLIKI 121
DB 541 AMKMETQIVATRAKGV-RLIYKEDGYLQAGATLIDI 575

RESULT 30
US-60-360-039-14608
Sequence 14608, Application US/60360039

GENERAL INFORMATION:

APPLICANT: Cao, Yongwei

APPLICANT: Chen, Xianfeng

APPLICANT: Goldman, Barry S.

APPLICANT: Hinkle, Gregory J.

APPLICANT: Slater, Steven C.

TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF

FILE REFERENCE: 38-10(52052)A

CURRENT APPLICATION NUMBER: US/60/360,039

CURRENT FILING DATE: 2002-02-21

NUMBER OF SEQ ID NOS: 47374

SEQ ID NO 14608

LENGTH: 575

TYPE: PRT

ORGANISM: Agrobacterium tumefaciens

US-60-360-039-14608

Matches 42; Conservative 17; Mismatches 28; Indels 9; Gaps 3;

OY 26 MGTILFGGTTGAPAPAGAGAGAGGEGTIPAPLAGTIVSKILYKEDGTVKAGOTVLYLE 85
DB 489 LGTV--SGGNASAPSAV-----EKKEGEMTAPVSGTLQSFVKDGETVSEGDLLAYME 540

OY 86 AMKMETEINAPTDGKVEKVLKERDAVVOGGGLIKI 121
DB 541 AMKMETQIVATRAKGV-RLIYKEDGYLQAGATLIDI 575

Search completed: May 1, 2003, 07:58:36
Job time: 178.469 secs

Query Match 28.7%; Score 176.5; DB 27; Length 575;
Best Local Similarity 43.8%; Pred. No. 4e-08;

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 1, 2003, 07:52:52 ; Search time 234.469 Seconds

(without alignments)
44.897 Million cell updates/sec

Title: US-09-987-485-1

Perfect score: 615
Sequence: 1 MKLKVYNGTAYDVVDVK.....KVLKERDAVQGGGLIKIG 122

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328255 seqs, 86286685 residues

Total number of hits satisfying chosen parameters: 328255

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database :

Published Applications AA: *
1: /cgn2_6/ptodata/2/pubppaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/2/pubppaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubppaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubppaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubppaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubppaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubppaa/PCTUS_PUBCOMB.pep:*
8: /cgn2_6/ptodata/2/pubppaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubppaa/US09_NEW_PUB.pep:*
10: /cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubppaa/US10_NEW_PUB.pep:*
12: /cgn2_6/ptodata/2/pubppaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/2/pubppaa/US60_NEW_PUB.pep:*
14: /cgn2_6/ptodata/2/pubppaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	615	100.0	122	10	US-09-987-485-1
2	342	55.6	70	10	US-09-987-485-2
3	173.5	28.2	134	9	US-10-169-048-16
4	170.5	27.7	591	9	US-09-738-626-6940
5	162.5	26.4	1157	9	US-09-974-973-2
6	162.5	26.4	1157	9	US-09-974-973-4
7	159.5	25.9	1140	9	US-09-974-973-19
8	159.5	25.9	1140	9	US-09-738-626-4265
9	159.5	25.9	1140	9	US-10-045-072-2
10	157	25.5	161	10	US-09-815-242-13363
11	157	25.5	161	10	US-09-815-242-13365
12	156	25.4	1073	10	US-09-815-242-12361
13	156	25.4	1147	10	US-09-815-242-5468
14	153	24.9	725	9	US-10-160-501-17
15	150	24.4	108	9	US-10-051-643-166
16	150	24.4	108	9	US-09-880-505-166
17	150	24.4	243	9	US-10-051-643-199
18	144	23.4	1142	9	US-09-815-242-10806
19	139.5	22.7	160	10	US-09-815-242-4962

20	139.5	22.7	162	10	US-09-815-242-10922	Sequence 10922, A
21	136.5	22.2	158	10	US-09-767-479-4	Sequence 4, Appl1
22	130	21.1	630	10	US-09-815-242-10036	Sequence 10036, A
23	129.5	21.1	182	10	US-09-767-479-2	Sequence 2, Appl1
24	120	19.5	154	10	US-09-815-242-12581	Sequence 12581, A
25	117	19.0	156	10	US-09-815-242-12062	Sequence 12062, A
26	115	18.7	71	10	US-09-791-171-48	Sequence 48, Appl
27	114.5	18.6	675	9	US-09-738-626-5917	Sequence 1159, A
28	112	18.2	155	10	US-09-815-242-11159	Sequence 16, Appl
29	107.5	17.5	402	9	US-10-081-051-16	Sequence 10329, A
30	103.5	16.8	156	10	US-09-815-242-13886	Sequence 11322, A
31	102.5	16.7	156	10	US-09-815-242-10329	Sequence 13886, A
32	101.5	16.5	156	10	US-09-815-242-11322	Sequence 10329, A
33	97.5	15.9	71	9	US-09-738-626-1965	Sequence 4965, Ap
34	95	15.4	2910	-9	US-10-124-800-2	Sequence 2, Appl1
35	90	14.6	229	10	US-09-815-242-5017	Sequence 5017, Ap
36	90	14.6	229	10	US-09-815-242-10613	Sequence 10613, A
37	89.5	14.6	539	10	US-09-815-242-10497	Sequence 10497, A
38	88	14.3	430	10	US-09-815-242-5802	Sequence 5802, Ap
39	88	14.3	430	10	US-09-815-242-12935	Sequence 12935, A
40	87	14.1	50	10	US-09-791-171-129	Sequence 129, App
41	83.5	13.6	229	10	US-09-815-242-13397	Sequence 13397, A
42	82.5	13.4	355	10	US-09-741-669-435	Sequence 435, App
43	80	13.0	50	10	US-09-791-171-81	Sequence 81, Appl
44	80	13.0	50	10	US-09-791-171-127	Sequence 127, Appl
45	78.5	12.8	380	10	US-09-823-240-10	Sequence 10, Appl
46	78	12.7	1657	9	US-10-123-155-433	Sequence 433, App
47	77.5	12.6	231	10	US-09-815-242-11998	Sequence 11998, A
48	77.5	12.6	234	10	US-09-815-242-10435	Sequence 10435, A
49	77.5	12.6	409	10	US-09-815-242-11270	Sequence 11270, A
50	77.5	12.6	2027	9	US-10-123-155-175	Sequence 175, App
51	76.5	12.4	424	10	US-09-815-242-12587	Sequence 12587, A
52	76.5	12.4	425	10	US-09-815-242-5485	Sequence 5485, Ap
53	76.5	12.4	819	9	US-09-820-843A-62	Sequence 62, Appl
54	76	12.4	2265	9	US-10-184-644-607	Sequence 607, App
55	76	12.4	2265	9	US-10-184-644-607	Sequence 607, App
56	75.5	12.3	196	9	US-10-059-261-69	Sequence 69, Appl
57	75.5	12.3	196	9	US-10-059-261-177	Sequence 177, App
58	75.5	12.3	625	10	US-09-854-731-18	Sequence 18, Appl
59	74.5	12.1	328	8	US-10-091-038-4	Sequence 39, Appl
60	74.5	12.1	329	8	US-08-945-028-9	Sequence 4, Appl1
61	74.5	12.1	424	10	US-09-835-684-9	Sequence 9, Appl1
62	74.5	12.1	424	10	US-09-880-371-9	Sequence 9, Appl1
63	74.5	12.1	424	10	US-09-879-248-14	Sequence 14, Appl
64	74.5	12.1	18636	9	US-10-073-912-17	Sequence 17, Appl
65	74	12.0	301	8	US-08-902-572-14	Sequence 14, Appl
66	74	12.0	301	10	US-10-161-499-12	Sequence 12, Appl
67	74	12.0	301	10	US-09-862-179A-41	Sequence 41, Appl
68	74	12.0	623	10	US-09-854-731-4	Sequence 4, Appl1
69	74	12.0	3069	9	US-09-712-363-246	Sequence 246, App
70	73.5	12.0	529	9	US-09-975-719-321	Sequence 321, App
71	73.5	12.0	2249	9	US-10-184-644-273	Sequence 273, App
72	73.5	12.0	2249	9	US-10-184-644-273	Sequence 273, App
73	73	11.9	413	10	US-09-815-242-13998	Sequence 13998, A
74	73	11.9	892	9	US-09-952-267-5	Sequence 5, Appl1
75	73	11.9	1079	9	US-09-820-843A-20	Sequence 20, Appl
76	73	11.9	1670	9	US-10-123-155-325	Sequence 325, App
77	73	11.9	2764	9	US-10-184-644-117	Sequence 117, App
78	73	11.9	2764	9	US-10-184-644-117	Sequence 117, App
79	73	11.9	2845	9	US-10-123-155-207	Sequence 207, App
80	72.5	11.8	234	10	US-09-815-242-14046	Sequence 14046, A
81	72.5	11.8	407	9	US-10-078-107-2	Sequence 2, Appl1
82	72.5	11.8	407	9	US-10-077-751-2	Sequence 2, Appl1
83	72.5	11.8	407	10	US-09-784-626-4	Sequence 4, Appl1
84	72.5	11.8	482	10	US-09-828-644-67	Sequence 473, Ap
85	72.5	11.8	595	10	US-10-184-644-261	Sequence 67, Appl
86	72.5	11.8	3038	9	US-10-184-644-261	Sequence 261, App
87	72.5	11.8	3038	9	US-10-123-155-409	Sequence 409, App
88	72.5	11.8	3617	9	US-09-837-669A-21	Sequence 21, Appl
89	72	11.7	166	10	US-09-841-821A-21	Sequence 21, Appl
90	72	11.7	200	10	US-09-764-864-953	Sequence 953, App
91	72	11.7	200	10	US-09-764-864-953	Sequence 953, App
92	72	11.7	355	9	US-09-712-363-161	Sequence 161, App

93 72 11.7 739 9 US-09-874-162A-5 Sequence 5, Appl
94 72 11.7 2171 9 US-10-184-644-515 Sequence 515, App
95 72 11.7 2171 9 US-10-184-634-515 Sequence 515, App
96 72 11.7 2285 9 US-10-184-644-111 Sequence 111, App
97 72 11.7 2285 9 US-10-184-634-111 Sequence 111, App
98 72 11.7 2368 9 US-10-123-155-423 Sequence 423, App
99 72 11.7 3305 9 US-10-184-644-87 Sequence 87, Appl
100 72 11.7 3305 9 US-10-184-634-87 Sequence 87, Appl

ALIGNMENTS

RESULT 1
US-09-987-485-1
; Sequence 1, Application US/09987485
; Patent No. US20020142355A1
; GENERAL INFORMATION:
; APPLICANT: Barry, Michael
; APPLICANT: Parrott, Michael
; TITLE OF INVENTION: Methods for the In Vivo Biotin Labeling of Polypeptides
; FILE REFERENCE: 15987/268653
; CURRENT APPLICATION NUMBER: US/09/987,485
; CURRENT FILING DATE: 2001-11-14
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 1
; LENGTH: 122
; TYPE: PRF
; ORGANISM: Escherichia coli
US-09-987-485-1

Query Match 100.0%; Score 615; DB 10; Length 122;
Best Local Similarity 100.0%; Pred. No. 1.9e-49;
Matches 122; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKKLVNCGAYDVVDVDSKSHENPMGTILFGGCTGAPAPAGAGAGAGGEPAPL 60
DB 1 MKKLVNCGAYDVVDVDSKSHENPMGTILFGGCTGAPAPAGAGAGAGGEPAPL 60
QY 61 AGVSKILVKEGDTKAGOTVLEAMKMETEINAPTDGKVEKVLKERDAVGGGLIK 120
DB 61 AGVSKILVKEGDTKAGOTVLEAMKMETEINAPTDGKVEKVLKERDAVGGGLIK 120
QY 121 IG 122
DB 121 IG 122

RESULT 2
US-09-987-485-2
; Sequence 2, Application US/09987485
; Patent No. US20020142355A1
; GENERAL INFORMATION:
; APPLICANT: Barry, Michael
; APPLICANT: Parrott, Michael
; TITLE OF INVENTION: Methods for the In Vivo Biotin Labeling of Polypeptides
; FILE REFERENCE: 15987/268653
; CURRENT APPLICATION NUMBER: US/09/987,485
; CURRENT FILING DATE: 2001-11-14
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 2
; LENGTH: 70
; TYPE: PRF
; ORGANISM: Escherichia coli
US-09-987-485-2

Query Match 55.6%; Score 342; DB 10; Length 70;
Best Local Similarity 100.0%; Pred. No. 1.1e-24;
Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 53 EGEIPAPLAGTVSKILVKEGDTKAGOTVLEAMKMETEINAPTDGKVEKVLKERDAV 112

DB 1 EGEIPAPLAGTVSKILVKEGDTKAGOTVLEAMKMETEINAPTDGKVEKVLKERDAV 60
QY 113 OGGGGLIKIG 122
DB 61 OGGGGLIKIG 70

RESULT 3
US-10-169-048-16
; Sequence 16, Application US/10169048
; Publication No. US20030072769A1
; GENERAL INFORMATION:
; APPLICANT: Clarke, Edna Elizabeth
; APPLICANT: Zhou, Liqing
; APPLICANT: Shea, Jacqueline Elizabeth
; APPLICANT: Feldman, Robert Graham
; APPLICANT: Holden, David William
; TITLE OF INVENTION: Streptococcus Pyogenes Virulence Genes and Proteins and Their
; FILE REFERENCE: GJE-97
; CURRENT APPLICATION NUMBER: US/10/169,048
; CURRENT FILING DATE: 2002-06-24
; PRIOR APPLICATION NUMBER: PCT/GB00/04997
; PRIOR FILING DATE: 2000-12-22
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 16
; LENGTH: 134
; TYPE: PRF
; ORGANISM: Streptococcus pyogenes
US-10-169-048-16

Query Match 28.2%; Score 173.5; DB 9; Length 134;
Best Local Similarity 32.4%; Pred. No. 6.2e-09;
Matches 47; Conservative 15; Mismatches 42; Indels 41; Gaps 2;

QY 2 KIKVTVNGAYDVVDVDSKSHENPMGTILFGGCTGAPAPAGAGAGAGAGGEPAPL 46
DB 5 KFKITIDGKREYILEME-----EIGAPAAAPAPPISTPVVPTESP 47
QY 47 -----GAGKAGEGEPAPLAGTVSKILVKEGDTKAGOTVLEAMKMETEINAPTDG 97
DB 48 QVEEAPAPQVPAAGADAIIPSPPTITILVAVGQVTEENPDLITLLEAMKMETEINAP 107
QY 98 DGRKVEKVLKERDAVGGGGLIKIG 122
DB 108 AGTITAIHVGPQGVVNPBGDLITIG 132

RESULT 4
US-09-738-626-6940
; Sequence 6940, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHITAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAOKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988

```
5
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 6940
; LENGTH: 591
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-6940

Query Match
Best Local Similarity 38.7%; Pred. No. 6.3e-08;
Matches 48; Conservative 16; Mismatches 45; Indels 15; Gaps 3;

OY 2 KLTATNGATAVDVDDKSHENPMGTLFEGGTCGAPAPA---AGAGACKAGEGEP 57
Db 478 KYVEINERREVALPGD-----LALGTAQPKKAKRRAGAKAGVSGDA-VA 526

OY 58 APLGTAKIIVKEDGYKAGOTVLYLEAMKMEFINAPTOGKVKVLYKENDAVOGGOG 117
Db 527 APMQSTVTKVNEGAEVNEGDTVYVLEAMKMEFINAPKAKSGTGTGLVAAGEGVNKGIV 586

OY 118 LIKI 121
Db 587 LLEI 590

RESULT 5
US-09-974-973-2
; Sequence 2, Application US/09974973
; Patent No. US20020177202A1
; GENERAL INFORMATION:
; APPLICANT: Hanke, Paul D.
; TITLE OF INVENTION: Feedback-Resistant Pyruvate Carboxylase Gene from Corynebacterium
; FILE REFERENCE: 1533.1230001/MAC/RGM
; CURRENT APPLICATION NUMBER: US/09/974.973
; PRIOR FILING DATE: 2001-10-21
; PRIOR APPLICATION NUMBER: US 60/239,913
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 1157
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-974-973-2

Query Match
Best Local Similarity 26.4%; Score 162.5; DB 9; Length 1157;
Matches 33; Conservative 14; Mismatches 26; Indels 1; Gaps 1;

OY 48 AGKAGEGEPAPLAGTAVSKIIYKEDGYKAGOTVLYLEAMKMEFINAPTDGKVKVLYK 107
Db 1084 ADSNKGHVAAPFAGVVT-VTVAEDEYKAGDAVAIIEMKMEATITASVDGKIERVYVP 1142

OY 108 ERDAVOGGGOLIKI 121
Db 1143 AATKVEGGDLIVV 1156

RESULT 6
US-09-974-973-4
; Sequence 4, Application US/09974973
; Patent No. US20020177202A1
; GENERAL INFORMATION:
; APPLICANT: Hanke, Paul D.
; TITLE OF INVENTION: Feedback-Resistant Pyruvate Carboxylase Gene from Corynebacterium
; FILE REFERENCE: 1533.1230001/MAC/RGM
; CURRENT APPLICATION NUMBER: US/09/974.973
; PRIOR FILING DATE: 2001-10-21
; PRIOR APPLICATION NUMBER: US 60/239,913
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 4
; LENGTH: 1157
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-974-973-4

Query Match
Best Local Similarity 26.4%; Score 162.5; DB 9; Length 1157;
Matches 33; Conservative 14; Mismatches 26; Indels 1; Gaps 1;

OY 48 AGKAGEGEPAPLAGTAVSKIIYKEDGYKAGOTVLYLEAMKMEFINAPTDGKVKVLYK 107
Db 1084 ADSNKGHVAAPFAGVVT-VTVAEDEYKAGDAVAIIEMKMEATITASVDGKIERVYVP 1142

OY 108 ERDAVOGGGOLIKI 121
Db 1143 AATKVEGGDLIVV 1156

RESULT 7
US-09-974-973-19
; Sequence 19, Application US/09974973
; Patent No. US20020177202A1
; GENERAL INFORMATION:
; APPLICANT: Hanke, Paul D.
; TITLE OF INVENTION: Feedback-Resistant Pyruvate Carboxylase Gene from Corynebacter
; FILE REFERENCE: 1533.1230001/MAC/RGM
; CURRENT APPLICATION NUMBER: US/09/974.973
; PRIOR FILING DATE: 2001-10-21
; PRIOR APPLICATION NUMBER: US 60/239,913
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 19
; LENGTH: 1140
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-974-973-19

Query Match
Best Local Similarity 25.9%; Score 159.5; DB 9; Length 1140;
Matches 32; Conservative 15; Mismatches 26; Indels 1; Gaps 1;

OY 48 AGKAGEGEPAPLAGTAVSKIIYKEDGYKAGOTVLYLEAMKMEFINAPTDGKVKVLYK 107
Db 1067 ADSNKGHVAAPFAGVVT-VTVAEDEYKAGDAVAIIEMKMEATITASVDGKIERVYVP 1125

OY 108 ERDAVOGGGOLIKI 121
Db 1126 AATKVEGGDLIVV 1139

RESULT 8
US-09-738-626-4265
; Sequence 4265, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAMA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAOKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738.626
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
```

```

? PRIOR APPLICATION NUMBER: JP 00/159166
? PRIOR FILING DATE: 2000-04-07
? PRIOR APPLICATION NUMBER: JP 00/280988
? PRIOR FILING DATE: 2000-08-03
? NUMBER OF SEQ ID NOS: 7059
? SOFTWARE: Patentln ver. 3.0
? SEQ ID NO 4265
? LENGTH: 1140
? TYPE: prt
? ORGANISM: Corynebacterium glutamicum
JS-09-736-626-4265

```

Query Match	25.98;	Score 159.5;	DB 9;	Length 1140;
Best Local Similarity	43.20;	Pred. No. 1.4e-06;		
Matches 32;	Conservative 15;	Mismatches 26;	Indels 1;	Gaps 1;

```

QY 48 AGRAGEGELPAPLACTVSKILVKEGTVKACQTVLVEAMKMETEINAPTDGKVEKVLK 107
      | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Db 1067 ADSSNNKGHVAAPFAGVVT-VTVAEGDEVKAGDAVAIIEMKMEATITFASVDGKIDRVVVP 1129

```

QY	108	ERDAVGGGGLIKI	121
		1:11 :: :	
Db	1126	AATKVEGGDLIVV	1139

RESULT 9
US-10-045-072-2
; Sequence 2, Application US/10045072
; Publication No. US20030027305A1
; GENERAL INFORMATION.

Query Match	25.98	Score 159.5	DB 9	Length 1140
Best Local Similarity	43.28	Pred. No. 1.4e-06		
Matches 32	Conservative 15	Mismatches 26		
				Gaps 1

Oy 48 AGAGEGEGLPAPLAGTSKLVKEDDTVAKGQTVLVEAMKKEITEINAPTDGVEKVLVK 107
 | : : : | | | : : | | | | | : : | | | | | | : : : : | : |
 Db 1067 ADSNNKGHVAAPFAGVVT -TVAAEEDEVKAGDAVAIIEMKKEATITLASVDGIDRVYVP 1125

```
QY      108 ERDAVQGGGGLIKI 121
          ||| ::
Db      1126 AATKVEGGDLIVV 1139
```

RESULT 10
US-09-815-242-13363
; Sequence 13363, Application US/09815242
; Patent No. US20020061569A1
; Grant No. 6480000

```

1  APPLICANT: Yamamoto, Robert T.
2  APPLICANT: Xu, H. Howard
3  TITLE OF INVENTION: Identification of Essential Genes in
4  TITLE OF INVENTION: Prokaryotes
5  FILE REFERENCE: ELITRA.011A
6  CURRENT APPLICATION NUMBER: US/09/815,242
7  CURRENT FILING DATE: 2001-03-21
8  PRIOR APPLICATION NUMBER: 60/191,078
9  PRIOR FILING DATE: 2000-03-21
10 PRIOR APPLICATION NUMBER: 60/206,848
11 PRIOR FILING DATE: 2000-05-23
12 PRIOR APPLICATION NUMBER: 60/207,727
13 PRIOR FILING DATE: 2000-05-26
14 PRIOR APPLICATION NUMBER: 60/242,578
15 PRIOR FILING DATE: 2000-10-23
16 PRIOR APPLICATION NUMBER: 60/253,625
17 PRIOR FILING DATE: 2000-11-27
18 PRIOR APPLICATION NUMBER: 60/257,921
19 PRIOR FILING DATE: 2000-12-22
20 PRIOR APPLICATION NUMBER: 60/269,308
21 PRIOR FILING DATE: 2001-02-16
22 NUMBER OF SEQ ID NOS: 14110
23 SOFTWARE: PastedQ for Windows Version 4.0
24 SEQ ID NO 1363
25 LENGTH: 161
26 TYPE: PRT
27 ORGANISM: Streptococcus pneumoniae
28 US-09-815-242-13363

```

Query Match	25.58;	Score 157;	DB 10;	Length 161;
Best Local Similarity	41.88;	Pred. No. 2.5e-07;		
Matches	38;	Conservative	15;	Mismatches 26;
			Indels	12;
			Gaps	3

```

Oy 39 PAPAGGAGGAKAGEGE-IPAPIAGTV-----SKIVKEGDTVAKGQFVILEAMKME 90
    ||| | | | : ||| | | | | | | | | | | | | | | | | | | | | | |
Db 74 PAPAE---ASVATEGMLVESPLGVVLLAAGPDKPAFVTVGDSYKKGQTLVILEAMKVM 129

```

```

Oy  91 TEINAPTDGKVEKVLKERRDAVOGGGLIKI 121
      || || || : || : | : || : ||
Db 130 NEIPAPKDGVTTEILVSNEMEVEFGKGLVRI 160

```

RESULT 11
US-09-815-242-13615
Sequence 13615, Application US/09815242
Patent No. US2002061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsem, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Travick, John D.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes In
FILE REFERENCE: ELTRA 011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308


```
;; TITLE OF INVENTION: MOLECULES AND USES THEREFOR
;; FILE REFERENCE: MNI-250
;; CURRENT APPLICATION NUMBER: US/10/160,501
;; CURRENT FILING DATE: 2002-05-30
;; PRIOR APPLICATION NUMBER: US 09/838,573
;; PRIOR FILING DATE: 2002-04-18
;; PRIOR APPLICATION NUMBER: US 60/197,747
;; PRIOR FILING DATE: 2000-04-18
;; PRIOR APPLICATION NUMBER: US 09/870,133
;; PRIOR FILING DATE: 2001-05-29
;; PRIOR APPLICATION NUMBER: US 60/207,649
;; PRIOR FILING DATE: 2000-05-26
;; PRIOR APPLICATION NUMBER: US 09/870,130
;; PRIOR FILING DATE: 2001-05-29
;; PRIOR APPLICATION NUMBER: US 60/207,640
;; PRIOR FILING DATE: 2000-05-26
;; PRIOR APPLICATION NUMBER: US 09/862,535
;; PRIOR FILING DATE: 2001-05-21
;; PRIOR APPLICATION NUMBER: US 60/205,961
;; PRIOR FILING DATE: 2000-05-19
;; PRIOR APPLICATION NUMBER: US 09/870,383
;; PRIOR FILING DATE: 2001-05-29
;; PRIOR APPLICATION NUMBER: US 60/207,506
;; PRIOR FILING DATE: 2000-05-26
;; PRIOR APPLICATION NUMBER: US 09/860,821
;; PRIOR FILING DATE: 2001-05-18
;; PRIOR APPLICATION NUMBER: US 60/205,449
;; PRIOR FILING DATE: 2000-05-19
;; PRIOR APPLICATION NUMBER: US 09/870,110
;; PRIOR FILING DATE: 2001-05-29
;; PRIOR APPLICATION NUMBER: US 60/207,650
;; PRIOR FILING DATE: 2000-05-26
;; PRIOR APPLICATION NUMBER: US 09/907,509
;; PRIOR FILING DATE: 2001-07-16
;; PRIOR APPLICATION NUMBER: US 60/218,385
;; PRIOR FILING DATE: 2000-07-14
;; PRIOR APPLICATION NUMBER: US 09/945,327
;; PRIOR FILING DATE: 2001-08-31
;; PRIOR APPLICATION NUMBER: US 60/229,425
;; PRIOR FILING DATE: 2000-08-31
;; PRIOR APPLICATION NUMBER: US 60/318,581
;; PRIOR FILING DATE: 2001-09-10
;; NUMBER OF SEQ ID NOS: 41
;; SOFTWARE: FastSeq Version 4.0
;; SEQ ID NO 17
;; LENGTH: 725
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-160-501-17

Query Match          24.9%; Score 153; DB 9; Length 725;
Best Local Similarity 37.3%; Pred. No. 3.2e-06;
Matches 41; Conservative 17; Mismatches 42; Indels 10; Gaps 3;

QY 3 LKVTNAGTAVYDVVDKSHENPMTILFSGGTG---AAPAAGAGAGAKAGEELPA 58
DB 598 LKCSVNGVAGSAKLIT---LEN---TILFSEKSGSIEIDIPVPPYLSVSSQETQGGPLA 651
QY 59 PLACTVSKILVKEGDTYKAGQTVLLEAMKMETEINAPTDGKVEKVLVKE 108
DB 652 PMETIEVFYKAGDKYKAGDSLWMTAMKMEHTIKSPKDGTVKVVYRE 701

RESULT 15
US-10-051-643-166
; Sequence 166, Application US/10051643
; Publication No. US20020197265A1
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Tan, Paul L. J.
; TITLE OF INVENTION: Methods and Compounds for the Treatment
; TITLE OF INVENTION: of Immunologically-Mediated Diseases of the Respiratory
; TITLE OF INVENTION: System using Mycobacterium Vaccae
```

```
;; FILE REFERENCE: 11000.1008c2
;; CURRENT APPLICATION NUMBER: US/10/051,643
;; CURRENT FILING DATE: 2002-01-18
;; PRIOR APPLICATION NUMBER: US09/156,181
;; PRIOR FILING DATE: 1998-09-17
;; PRIOR APPLICATION NUMBER: US 08/996,624
;; PRIOR FILING DATE: 1997-12-23
;; NUMBER OF SEQ ID NOS: 208
;; SOFTWARE: FastSeq for Windows Version 3.0
;; SEQ ID NO 166
;; LENGTH: 108
;; TYPE: PRT
;; ORGANISM: Mycobacterium vaccae
US-10-051-643-166

Query Match          24.4%; Score 150; DB 9; Length 108;
Best Local Similarity 46.5%; Pred. No. 6.9e-07;
Matches 33; Conservative 10; Mismatches 28; Indels 0; Gaps 0;

QY 52 GEGEIPAPLAGTAVSKILVKEGDTYKAGQTVLLEAMKMETEINAPTDGKVEKVLKERDA 111
DB 33 GDERVDAPEFASVWKVDVAVGDRVAGQPLALEAMKMETEVLRAPADGVYVTLIVSAGHL 92
QY 112 VGGGGLIKIG 122
DB 93 VDPGTPLVVVG 103

RESULT 16
US-09-880-505-166
; Sequence 166, Application US/09880505
; Publication No. US20030007976A1
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Prestidge, Ross
; TITLE OF INVENTION: Methods and Compounds for the Treatment
; TITLE OF INVENTION: of Immunologically-Mediated Skin Disorders
; FILE REFERENCE: 11000.1007c2
; CURRENT APPLICATION NUMBER: US/09/880,505
; CURRENT FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 09/324,542
; PRIOR FILING DATE: 1999-06-02
; PRIOR APPLICATION NUMBER: US 08/997,080
; PRIOR FILING DATE: 1997-12-23
; NUMBER OF SEQ ID NOS: 194
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 166
; LENGTH: 108
; TYPE: PRT
; ORGANISM: Mycobacterium vaccae
US-09-880-505-166

Query Match          24.4%; Score 150; DB 9; Length 108;
Best Local Similarity 46.5%; Pred. No. 6.9e-07;
Matches 33; Conservative 10; Mismatches 28; Indels 0; Gaps 0;

QY 52 GEGEIPAPLAGTAVSKILVKEGDTYKAGQTVLLEAMKMETEINAPTDGKVEKVLKERDA 111
DB 33 GDERVDAPEFASVWKVDVAVGDRVAGQPLALEAMKMETEVLRAPADGVYVTLIVSAGHL 92
QY 112 VGGGGLIKIG 122
DB 93 VDPGTPLVVVG 103

RESULT 17
US-10-051-643-199
; Sequence 199, Application US/10051643
; Publication No. US20020197265A1
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Tan, Paul L. J.
```


TITLE OF INVENTION: Methods and Compounds for the Treatment
; TITLE OF INVENTION: of Immunologically-Mediated Diseases of the Respiratory
; FILE REFERENCE: 11000.10082
; CURRENT APPLICATION NUMBER: US/10/051,643
; PRIOR FILING DATE: 2002-01-18
; PRIOR APPLICATION NUMBER: US09/156,181
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: US 08/996,624
; PRIOR FILING DATE: 1997-12-23
; NUMBER OF SEQ ID NOS: 208
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 199
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Mycobacterium vaccae
US-10-051-643-199

Query Match 24.4%; Score 150; DB 9; Length 243;
Best Local Similarity 46.5%; Pred. No. 1.7e-06;
Matches 33; Conservative 10; Mismatches 28; Indels 0; Gaps 0;

QY 52 GGGELPAGTAVSKILYKGGDTVAGQTVLLEAMKMETINAPTDGKVEKVRKEND 111
DB 168 GDERVDAPFASVWVAVAGDGVVAGGPLLALEMKMETIVRAPDGVQILVSAGHL 227

QY 112 VGGGGLIKIG 122
DB 228 VDPGTPPLVVG 238

RESULT 18
US-09-815-242-10806
; Sequence 10806, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10806
; LENGTH: 1142
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-815-242-10806

Query Match 23.4%; Score 144; DB 10; Length 1142;
Best Local Similarity 43.3%; Pred. No. 3.6e-05;

Matches 29; Conservative 17; Mismatches 21; Indels 0; Gaps 0;

QY 55 EIPAPLAGTAVSKILYKGGDTVAGQTVLLEAMKMETINAPTDGKVEKVRKENDAVOG 114
DB 1073 QIGATMSGVLAQVLYKRDGKVEKGPLLITLTKMKMETITLFACTVTHIYEEBAISS 1132

QY 115 GGGELIKI 121
DB 1133 GDLLEEV 1139

RESULT 19
US-09-815-242-4962
; Sequence 4962, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4962
; LENGTH: 160
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-815-242-4962

Query Match 22.7%; Score 139.5; DB 10; Length 160;
Best Local Similarity 40.5%; Pred. No. 9.8e-06;
Matches 30; Conservative 12; Mismatches 25; Indels 7; Gaps 1;

QY 55 EIPAPLAGTAVSKILYKGGDTVAGQTVLLEAMKMETINAPTDGKVEKVRKEND 107
DB 87 ELTSPVIGIVLQAPADKNEFKVGDYTKGVGVCIKEMKMLNETLTFVDSVITLILVN 146

QY 108 EENDAVGGGGLIKI 121
DB 147 NEDVVEFGGPIFRV 160

RESULT 20
US-09-815-242-10922
; Sequence 10922, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel

US-09-815-242-10036

Query Match 21.1%, Score 130; DB 10; Length 630;
Best Local Similarity 31.6%, Pred. No. 0.00035;
Matches 43; Conservative 20; Mismatches 53; Indels 20; Gaps 4;

QY 1 MKLAVTNGTAVDVV-DKSHENPMGTILFGGTGAPAPAGAGAGAGECEIPAP 59
DB 147 MEVPAPFAGTVEIKVNGDKYSTSLIMVEFVAGEAAPAAQEAAPAAA-----PAP 202
QY 60 LAGT-----VSKLVEGDTKAGOTVLEAMKMTETINAPTDKVEEVL 105
DB 203 AAGVKEVAVPDIGDEVEETEMVAVGKVAEBSLITVEGDKASMEVPAPAGVVEIK 262
QY 106 VKERDAVGGGGLIKI 121
DB 263 VNVGDKVKTG-SLIMI 277

RESULT 23

US-09-767-479-2
Sequence 2, Application US/09767479
Patent No. US20010036654A1

GENERAL INFORMATION:

APPLICANT: Haselkorn, Robert

Gornicki, Piotr

TITLE OF INVENTION: ACETYL-COA CARBOXYLASE COMPOSITIONS AND METHODS OF USE

NUMBER OF SEQUENCES: 29

CORRESPONDENCE ADDRESS:

ADDRESS: Arnold, White & Durkee

STREET: P.O. Box 4433

CITY: Houston

STATE: Texas

COUNTRY: United States of America

ZIP: 77210

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/767,479

FILING DATE: 22-Jan-2001

CLASSIFICATION: Unknown

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/468,793

FILING DATE: <Unknown>

APPLICATION NUMBER: US SN 07/956,700

FILING DATE: 02-OCT-1992

APPLICATION NUMBER: PCT/US93/09340

FILING DATE: 30-SEP-1993

ATTORNEY/AGENT INFORMATION:

NAME: Kitchell, Barbara S.

REGISTRATION NUMBER: 33,928

REFERENCE/DOCKET NUMBER: ARCD:152/KIT

TELECOMMUNICATION INFORMATION:

TELEPHONE: (512) 418-3000

TELEFAX: (713) 789-2679

TELEX: 79-0924

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 182 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 2:

US-09-767-479-2

QY 34 GTGAPAPAGAGAGAGAGE-----GEIPAPLAGTVSK-----ILVKEGDTVKAGOT 80
DB 79 GTSRADIAVAVSSGSPAKIILDKLAVASPMVGTFFRAAPGRAVVEVGDRIROGOT 138
QY 81 VLVEAMKMTETINAPTDKVEKVLVKERDAVGGGGLIKI 121
DB 139 VCIIEMAKLMEIEADVSGQVTEILVONGEPEVEYNQPIMRI 179

RESULT 24

US-09-815-242-12581

Sequence 12581, Application US/09815242

Patent No. US20020061569A1

GENERAL INFORMATION:

APPLICANT: Haselbeck, Robert

Ohlser, Karl L.

APPLICANT: Zyskind, Judith W.

APPLICANT: Wall, Daniel

APPLICANT: Trawick, John D.

APPLICANT: Carr, Grant J.

APPLICANT: Yamamoto, Robert T.

APPLICANT: Xu, H. Howard

TITLE OF INVENTION: Identification of Essential Genes in

FILE REFERENCE: ELITRA.011A

CURRENT APPLICATION NUMBER: US/09/815,242

CURRENT FILING DATE: 2001-03-21

PRIOR APPLICATION NUMBER: 60/191,078

PRIOR FILING DATE: 2000-03-21

PRIOR APPLICATION NUMBER: 60/206,848

PRIOR FILING DATE: 2000-05-23

PRIOR APPLICATION NUMBER: 60/207,727

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: 60/242,578

PRIOR FILING DATE: 2000-10-23

PRIOR APPLICATION NUMBER: 60/253,625

PRIOR FILING DATE: 2000-11-27

PRIOR APPLICATION NUMBER: 60/257,931

PRIOR FILING DATE: 2000-12-22

PRIOR APPLICATION NUMBER: 60/269,308

PRIOR FILING DATE: 2001-02-16

NUMBER OF SEQ ID NOS: 14110

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 12581

LENGTH: 154

TYPE: PRT

ORGANISM: Staphylococcus aureus

US-09-815-242-12581

Query Match 19.5%, Score 120; DB 10; Length 154;

Best Local Similarity 35.2%, Pred. No. 0.00058;

Matches 32; Conservative 13; Mismatches 32; Indels 14; Gaps 2;

QY 38 APAPAGAGAGAGAGEGIPAPLAGTVSK-----ILVKEGDTVKAGOTVLEAMKME 90
DB 70 AKPTSD-----NHKTIAPMVGTFFKSPSPDEAVVQGDVTSNFTVCIIEMAKLTF 122
QY 91 TEINAPTDKVEKVLVKERDAVGGGGLIKI 121
DB 123 NEIQAEISGEIIEILVEGDKVMEYGOPLKXV 153

RESULT 25

US-09-815-242-12062

Sequence 12062, Application US/09815242

Patent No. US20020061569A1

GENERAL INFORMATION:

APPLICANT: Haselbeck, Robert

Ohlser, Karl L.

APPLICANT: Zyskind, Judith W.

APPLICANT: Wall, Daniel

APPLICANT: Trawick, John D.

```

; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA-011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12062
; LENGTH: 156
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; US-09-815-242-12062

Query Match          19.0%; Score 117; DB 10; Length 156;
Best Local Similarity 32.0%; Pred. No. 0.0011;
Matches 32; Conservative 12; Mismatches 40; Indels 16; Gaps 2;

QY 38 AAPAAGAGAG-----KAGEEIPAPLAGT-----VSKILVKEGDFVKAQGV 81
DB 56 AAPAPAAAAPAAAASAPAAVKLGNVYRSPMGTFYRAASPSANVEVGGVKKGDIL 115
QY 82 LVLEAMKETEINAPTKYKVEKLVKERDAVGGGGLIKI 121
DB 116 CIVEAMKMMHIEVSGTIESILVENGQVPEFQPLEFTI 155

RESULT 26
; US-09-791-171-48
; Sequence 48; Application US/09/91171
; Patent No. US20020094336A1
; GENERAL INFORMATION:
; APPLICANT: ANDERSEN, Peter
; APPLICANT: NIELSEN, Rikke
; APPLICANT: OETTINGER, Thomas
; APPLICANT: RASMUSSEN, Peter Birk
; APPLICANT: ROSENKRANDS, Ida
; APPLICANT: WELDLINGH, Karin
; APPLICANT: FLORIO, Walter
; TITLE OF INVENTION: NUCLEIC ACIDS FRAGMENTS AND POLYPEPTIDE FRAGMENTS
; FILE REFERENCE: 670001-2002.1
; CURRENT APPLICATION NUMBER: US/09/791,171
; CURRENT FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 09/050,739
; PRIOR FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: 0376/97
; PRIOR FILING DATE: 1997-04-02
; PRIOR APPLICATION NUMBER: 1277/97
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: 60/044,624
; PRIOR FILING DATE: 1997-04-18
; PRIOR APPLICATION NUMBER: 60/070,488
; PRIOR FILING DATE: 1998-01-05
; NUMBER OF SEQ ID NOS: 173
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 48
```

```

; LENGTH: 71
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
; US-09-791-171-48

Query Match          18.7%; Score 115; DB 10; Length 71;
Best Local Similarity 37.7%; Pred. No. 0.00068;
Matches 23; Conservative 13; Mismatches 25; Indels 0; Gaps 0;

QY 55 EIPAPLAGTVSKILVKEGDFVKAQGVLVLEAMKETEINAPTKYKVEKLVKERDAVGG 114
DB 4 DVAREIYASVLEVVNNGDDIDKGDVYVLESKMEIPVLAAGTVSKYASVGVYQA 63
QY 115 G 115
DB 64 G 64

RESULT 27
; US-09-738-626-5917
; Sequence 5917; Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SETKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAOKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 5917
; LENGTH: 675
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
; US-09-738-626-5917

Query Match          18.6%; Score 114.5; DB 9; Length 675;
Best Local Similarity 27.3%; Pred. No. 0.0099;
Matches 35; Conservative 20; Mismatches 44; Indels 29; Gaps 3;

QY 15 DVDVDSHNPMTIIFGGGTGAPAP-----AGAGAGAGAGCEIP- 57
DB 78 DADETPANEAP-----ADEAPAPAEDEEPKKEAPKAPATGATDVEPME 128
QY 58 ---APLAGTVSKILVKEGDFVKAQGVLVLEAMKETEINAPTKYKVEKLVKERDAVGG 114
DB 129 LGEVTEGIIITQKLVAGVDEPLLEVSTDKVDTEIPSPAGTIVEILADEDDTVV 188
QY 115 GGLIKIRIG 122
DB 189 GAVIARIG 196

RESULT 28
; US-09-815-242-11159
; Sequence 11159; Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
```

```

: TYPE: PRT
: ORGANISM: Escherichia ruminantium (formerly Cowdria ruminantium)
US-10-081-051-16
Query Match
Best Local Similarity 17.5%; Score 107.5; DB 9; Length 402;
Matches 23; Conservative 15; Mismatches 27; Indels 5; Gaps 1;

OY 52 GCGEIPAPLAGTGVSKILVEGDPVTKAGQTVLLEAKMETEINAPTDGKVEKVERDA 111
DB 13 GCSILEAPI-----RVSKIGDSIKQGVLFITETDKTSLEIVSPVDGTVSVFLADEEI 67
OY 112 VGGGGGLIKI 121
DB 68 IERDQLCTI 77

RESULT 30
US-09-815-242-13886
: Sequence 13886, Application US/09815242
: Patent No. US20020061569A1
: GENERAL INFORMATION:
: APPLICANT: Haselbeck, Robert
: APPLICANT: Ohlsen, Karl L.
: APPLICANT: Zyskind, Judith W.
: APPLICANT: Wall, Daniel
: APPLICANT: Trawick, John D.
: APPLICANT: Carr, Grant J.
: APPLICANT: Yamamoto, Robert T.
: APPLICANT: Xu, H. Howard
: TITLE OF INVENTION: Identification of Essential Genes in
: FILE REFERENCE: ELITRA.011A
: CURRENT APPLICATION NUMBER: US/09/815,242
: PRIOR FILING DATE: 2001-03-21
: PRIOR APPLICATION NUMBER: 60/191,078
: PRIOR FILING DATE: 2000-03-21
: PRIOR APPLICATION NUMBER: 60/206,848
: PRIOR FILING DATE: 2000-05-23
: PRIOR APPLICATION NUMBER: 60/207,727
: PRIOR FILING DATE: 2000-05-26
: PRIOR APPLICATION NUMBER: 60/242,578
: PRIOR FILING DATE: 2000-10-23
: PRIOR APPLICATION NUMBER: 60/253,625
: PRIOR FILING DATE: 2000-11-27
: PRIOR APPLICATION NUMBER: 60/257,931
: PRIOR FILING DATE: 2000-12-22
: PRIOR APPLICATION NUMBER: 60/269,308
: PRIOR FILING DATE: 2001-02-16
: NUMBER OF SEQ ID NOS: 14110
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 11159
: LENGTH: 155
: TYPE: PRT
: ORGANISM: Haemophilus influenzae
US-09-815-242-11159
Query Match
Best Local Similarity 18.2%; Score 112; DB 10; Length 155;
Matches 30; Conservative 16; Mismatches 38; Indels 12; Gaps 2;

OY 38 APAPAGAGAGKAGEGE-----IPAPLAGTV-----SKILVEGDPVTKAGQTVLVE 85
DB 59 APAQVPAAPATTPAASDELGHLVSPMVGTYRSPSPKAFVEVGSGVKGDALCTYE 118
OY 86 AKMETEINAPTDGKVEKVERDAVGGGGLIKI 121
DB 119 AKMKNRIEADKAGVYKAILINDGNAVERDEPLVIY 154

RESULT 29
US-10-081-051-16
: Sequence 16, Application US/10081051
: Publication No. US2003004422A1
: GENERAL INFORMATION:
: APPLICANT: Barbet, Anthony F.
: APPLICANT: Whitliffe, William M.
: APPLICANT: Kamper, Sandra M.
: APPLICANT: Simbl, Bigboy H.
: APPLICANT: Ganta, Roman R.
: APPLICANT: Moreland, Annie L.
: APPLICANT: Mwangli, Duncan M.
: APPLICANT: Mcguire, Travis C.
: APPLICANT: Mahan, Suman M.
: TITLE OF INVENTION: Ehrlichia ruminantium Polypeptides, Antigens, Polynucleotides, an
: FILE REFERENCE: UF-299XC1
: CURRENT APPLICATION NUMBER: US/10/081,051
: PRIOR FILING DATE: 2002-02-20
: PRIOR APPLICATION NUMBER: US 60/269,944
: PRIOR FILING DATE: 2001-02-20
: NUMBER OF SEQ ID NOS: 117
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 16
: LENGTH: 402
```

```

: TYPE: PRT
: ORGANISM: Escherichia ruminantium (formerly Cowdria ruminantium)
US-10-081-051-16
Query Match
Best Local Similarity 17.5%; Score 107.5; DB 9; Length 402;
Matches 23; Conservative 15; Mismatches 27; Indels 5; Gaps 1;

OY 52 GCGEIPAPLAGTGVSKILVEGDPVTKAGQTVLLEAKMETEINAPTDGKVEKVERDA 111
DB 13 GCSILEAPI-----RVSKIGDSIKQGVLFITETDKTSLEIVSPVDGTVSVFLADEEI 67
OY 112 VGGGGGLIKI 121
DB 68 IERDQLCTI 77

RESULT 30
US-09-815-242-13886
: Sequence 13886, Application US/09815242
: Patent No. US20020061569A1
: GENERAL INFORMATION:
: APPLICANT: Haselbeck, Robert
: APPLICANT: Ohlsen, Karl L.
: APPLICANT: Zyskind, Judith W.
: APPLICANT: Wall, Daniel
: APPLICANT: Trawick, John D.
: APPLICANT: Carr, Grant J.
: APPLICANT: Yamamoto, Robert T.
: APPLICANT: Xu, H. Howard
: TITLE OF INVENTION: Identification of Essential Genes in
: FILE REFERENCE: ELITRA.011A
: CURRENT APPLICATION NUMBER: US/09/815,242
: PRIOR FILING DATE: 2001-03-21
: PRIOR APPLICATION NUMBER: 60/191,078
: PRIOR FILING DATE: 2000-03-21
: PRIOR APPLICATION NUMBER: 60/206,848
: PRIOR FILING DATE: 2000-05-23
: PRIOR APPLICATION NUMBER: 60/207,727
: PRIOR FILING DATE: 2000-05-26
: PRIOR APPLICATION NUMBER: 60/242,578
: PRIOR FILING DATE: 2000-10-23
: PRIOR APPLICATION NUMBER: 60/253,625
: PRIOR FILING DATE: 2000-11-27
: PRIOR APPLICATION NUMBER: 60/257,931
: PRIOR FILING DATE: 2000-12-22
: PRIOR APPLICATION NUMBER: 60/269,308
: PRIOR FILING DATE: 2001-02-16
: NUMBER OF SEQ ID NOS: 14110
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 13886
: LENGTH: 156
: TYPE: PRT
: ORGANISM: Salmonella typhi
US-09-815-242-13886
Query Match
Best Local Similarity 16.8%; Score 103.5; DB 10; Length 156;
Matches 28; Conservative 17; Mismatches 39; Indels 9; Gaps 2;

OY 38 APA-PAAGAGAGKAGEGEIPAPLAGTV-----SKILVEGDPVTKAGQTVLLEAK 88
DB 63 APAAPFAPAPAAEISGIVHSIPMVGTYRSPDAKAFIEVGQKVVNVDGLCTIYEA 122
OY 89 METEINAPTDGKVEKVERDAVGGGGLIKI 121
DB 123 MNQIEADKAGVYKAILVESQVPERDEPLVI 155

Search completed: May 1, 2003, 08:16:10
Job time : 235.469 secs
```

THIS PAGE BLANK (USPTO)

Gencore version 5.1.4_p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 1, 2003, 07:52:32 ; Search time 409.844 Seconds

(without alignments)
50.445 Million cell updates/sec

Title: US-09-987-485-1

Perfect score: 615

Sequence: 1 MKLKVTYNGTAYVDVDYDK.....KVLKRDVAGGGLIKIG 122

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 777667 seqs, 169465163 residues

Total number of hits satisfying chosen parameters: 777667

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : Pending Patents_AA_New:*

1: /cgn2_6/ptodata/2/paa/US06_NEW_COMB.pep:*

2: /cgn2_6/ptodata/2/paa/US07_NEW_COMB.pep:*

3: /cgn2_6/ptodata/2/paa/US08_NEW_COMB.pep:*

4: /cgn2_6/ptodata/2/paa/US09_NEW_COMB.pep:*

5: /cgn2_6/ptodata/2/paa/US10_NEW_COMB.pep:*

6: /cgn2_6/ptodata/2/paa/US10_NEW_COMB.pep:*

7: /cgn2_6/ptodata/2/paa/US60_NEW_COMB.pep:*

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	436	70.9	125 1 PCT-US02-32727-12317	Sequence 12317, A
2	436	70.9	125 5 US-09-978-825-12317	Sequence 12317, A
3	436	70.9	125 6 US-10-057-458-12317	Sequence 12317, A
4	204.5	33.3	589 6 US-10-369-493-10079	Sequence 10079, A
5	203.5	33.1	143 6 US-10-282-122A-48959	Sequence 48959, A
6	195	31.7	588 6 US-10-282-122A-60140	Sequence 60140, A
7	190.5	31.0	655 6 US-10-369-493-112	Sequence 112, App
8	187.5	30.5	599 6 US-10-282-122A-54468	Sequence 54468, A
9	185	30.1	597 6 US-10-282-122A-77079	Sequence 77079, A
10	184	29.9	571 6 US-10-369-493-1263	Sequence 1263, App
11	184	29.9	571 6 US-10-369-493-20329	Sequence 20329, A
12	180.5	29.3	1076 6 US-10-369-493-13831	Sequence 13831, A
13	180.5	28.7	574 6 US-10-369-493-13439	Sequence 13439, A
14	176.5	28.7	574 6 US-10-369-493-15017	Sequence 15017, A
15	176.5	28.7	575 6 US-10-369-493-11591	Sequence 11591, A
16	176.5	28.7	575 6 US-10-369-493-14608	Sequence 14608, A
17	175.5	28.5	147 5 US-09-134-000C-6221	Sequence 6221, App
18	175.5	28.5	147 5 US-09-134-000C-6221	Sequence 6221, App
19	175.5	28.5	602 6 US-10-282-122A-67750	Sequence 67750, A
20	175.5	28.5	610 6 US-10-369-493-18633	Sequence 18633, A
21	175	28.5	596 6 US-10-282-122A-61280	Sequence 61280, A
22	175	28.5	596 6 US-10-282-122A-61280	Sequence 61280, A
23	174	28.3	591 6 US-10-089-057-4	Sequence 46866, A
24	173.5	28.2	134 6 US-10-169-048-16	Sequence 16, App1
25	173.5	28.2	591 6 US-10-282-122A-75568	Sequence 75568, A
26	173.5	28.2	591 6 US-10-282-122A-76334	Sequence 76334, A

27	173	28.1	567	6	US-10-369-493-1053	Sequence 1053, App
28	172.5	28.0	602	6	US-10-282-122A-69598	Sequence 69598, A
29	170	27.6	1144	6	US-10-282-122A-52088	Sequence 52088, A
30	168.5	27.4	607	6	US-10-282-122A-43666	Sequence 43666, A
31	168.5	27.4	612	6	US-10-369-493-19134	Sequence 19134, A
32	168.5	27.4	612	6	US-10-419-128-19134	Sequence 19134, A
33	168.5	27.4	620	6	US-10-369-493-119	Sequence 119, App
34	165	26.8	1150	6	US-10-369-493-17338	Sequence 17338, A
35	164	26.7	654	6	US-10-369-493-16954	Sequence 16954, A
36	163	26.5	598	6	US-10-282-122A-63744	Sequence 63744, A
37	161.5	26.3	1147	6	US-10-369-493-11450	Sequence 11450, A
38	161.5	26.3	1151	6	US-10-369-493-14817	Sequence 14817, A
39	161.5	26.3	1151	6	US-10-369-493-15006	Sequence 15006, A
40	161	26.2	594	6	US-10-282-122A-53886	Sequence 53886, A
41	160.5	26.1	75	6	US-10-224-539A-7	Sequence 7, App11
42	160.5	26.1	75	6	US-10-370-097-133	Sequence 133, App
43	160.5	26.1	75	6	US-10-370-097-133	Sequence 133, App
44	160	26.0	91	6	US-10-282-122A-47349	Sequence 47349, A
45	158.5	25.8	593	6	US-10-282-122A-76375	Sequence 76375, A
46	158	25.7	1144	6	US-10-282-122A-52777	Sequence 52777, A
47	157.5	25.6	1141	6	US-10-282-122A-53885	Sequence 53885, A
48	157	25.5	1141	6	US-10-282-122A-73808	Sequence 73808, A
49	157	25.5	1146	6	US-10-282-122A-60488	Sequence 60488, A
50	156	25.4	611	6	US-10-282-122A-70190	Sequence 70190, A
51	156	25.4	989	5	US-09-950-084-3973	Sequence 3973, App
52	156	25.4	1150	6	US-10-282-122A-44391	Sequence 44391, A
53	154.5	25.1	1152	6	US-10-369-493-12027	Sequence 12027, A
54	153.5	25.0	162	6	US-10-282-122A-72184	Sequence 72184, A
55	153	24.9	156	6	US-10-282-122A-49663	Sequence 49663, A
56	153	24.9	725	6	US-10-224-539A-2	Sequence 2, App11
57	153	24.9	725	6	US-10-224-539A-9	Sequence 9, App11
58	153	24.9	725	6	US-10-370-959-128	Sequence 128, App
59	153	24.9	725	6	US-10-370-959-135	Sequence 135, App
60	153	24.9	725	6	US-10-370-959-135	Sequence 135, App
61	153	24.9	725	6	US-10-370-959-135	Sequence 135, App
62	153	24.9	725	6	US-60-453-135-10473	Sequence 10473, A
63	153	24.9	725	6	US-60-453-135-10473	Sequence 10473, A
64	152.5	24.8	602	6	US-10-282-122A-67292	Sequence 67292, A
65	152.5	24.8	901	6	US-10-369-493-10556	Sequence 10556, A
66	152	24.7	701	6	US-10-369-493-27999	Sequence 27999, A
67	152	24.7	701	6	US-10-419-128-77999	Sequence 77999, A
68	151	24.6	166	6	US-10-282-122A-74706	Sequence 74706, A
69	151	24.6	1142	6	US-10-282-122A-57942	Sequence 57942, A
70	150	24.4	108	5	US-09-880-505A-166	Sequence 166, App
71	150	24.4	190	6	US-10-417-884-5720	Sequence 5720, App
72	150	24.4	1143	6	US-10-282-122A-52917	Sequence 52917, A
73	149	24.2	611	6	US-10-282-122A-70702	Sequence 70702, A
74	149	24.2	1154	6	US-10-092-411A-3428	Sequence 3428, App
75	148.5	24.1	615	6	US-10-375-039-32	Sequence 32, App1
76	148.5	24.1	676	6	US-10-369-493-26143	Sequence 26143, A
77	148.5	24.1	676	6	US-10-419-128-26143	Sequence 26143, A
78	147	23.9	654	6	US-10-369-493-15564	Sequence 15564, A
79	147	23.9	654	6	US-10-369-493-15564	Sequence 15564, A
80	147	23.9	694	6	US-10-369-493-12310	Sequence 12310, A
81	147	23.9	1078	6	US-10-369-493-17039	Sequence 17039, A
82	146.5	23.8	1127	6	US-10-282-122A-62684	Sequence 62684, A
83	146.5	23.8	588	6	US-10-282-122A-73367	Sequence 73367, A
84	146.5	23.8	588	6	US-10-369-493-10569	Sequence 10569, A
85	145	23.6	631	6	US-10-282-122A-42558	Sequence 42558, A
86	144	23.4	1142	6	US-09-134-000C-5707	Sequence 5707, App
87	144	23.4	1163	5	US-09-134-000C-45491	Sequence 45491, A
88	144	23.4	1163	5	US-09-134-000C-45491	Sequence 45491, A
89	144	23.4	1167	6	US-10-369-493-22819	Sequence 22819, A
90	144	23.4	1185	6	US-10-369-493-42881	Sequence 42881, A
91	144	23.3	1137	6	US-10-369-493-18351	Sequence 18351, A
92	143.5	23.3	368	6	US-10-369-493-1203	Sequence 1203, App
93	141	22.9	611	6	US-10-282-122A-71785	Sequence 71785, A
94	141	22.9	162	6	US-10-282-122A-42441	Sequence 42441, A
95	139.5	22.7	168	6	US-09-134-000C-4552	Sequence 4552, App
96	139.5	22.7	168	6	US-09-134-000C-4552	Sequence 4552, App
97	139.5	22.6	671	6	US-10-369-493-7093	Sequence 7093, App
98	139	22.6	674	6	US-10-369-493-4338	Sequence 4338, App
99	139	22.6	674	6	US-10-369-493-4338	Sequence 4338, App

100 138.5 22.5 716 6 US-10-369-493-12942 Sequence 12942, A

ALIGNMENTS

RESULT 1
PCT-US02-32727-12317

Sequence 12317, Application PC/TUS0232727
GENERAL INFORMATION:
APPLICANT: Mitcham, Jennifer
APPLICANT: Skeiky, Yasir
APPLICANT: Persing, David
APPLICANT: Bhatia, Ajay
APPLICANT: Maisonneuve, Jean Francois
APPLICANT: Zhang, Yanni
APPLICANT: Wang, Siqing
APPLICANT: Jen, Shyian
APPLICANT: Lodes, Michael
APPLICANT: Benson, Darin
APPLICANT: Jones, Robert
APPLICANT: Carter, Darlick
APPLICANT: Barth, Brenda
APPLICANT: Douglass, John
TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acnes
FILE REFERENCE: 210121.514C1
CURRENT APPLICATION NUMBER: PCT/US02/32727
CURRENT FILING DATE: 2002-10-11
NUMBER OF SEQ ID NOS: 30992
SEQ ID NO 12317
LENGTH: 125
TYPE: PRT
ORGANISM: Propionibacterium acnes
PCT-US02-32727-12317

Query Match 70.9%; Score 436; DB 1; Length 125;
Best Local Similarity 73.8%; Pred. No. 1.6e-29;
Matches 90; Conservative 9; Mismatches 21; Indels 2; Gaps 2;

QY 1 MKLKVYNGTAYDVVDVYKSHENPMGTILFGGCTGAPAPAAAGAGAGAGGEGEIPAPL 60
DB 6 MKLKVYNDVAYDVVDVDTANAPAPAPILFGGAGG-PMKASGG-GAGKAGEGEVAPL 63
QY 61 AGTYSKILVKEGDTVRAGOTVIVLEAMKMETEINAPTDGKVEKYLKERDAVGGGGLIK 120
DB 64 AGTAAKILVABGDVAVKAGVLTLEAMKMETEINAPADGTVKGIIVAVGDVAVGGGGLVA 123

QY 121 IG 122
DB 124 LG 125

RESULT 2
US-09-978-825-12317
Sequence 12317, Application US/09978825
GENERAL INFORMATION:
APPLICANT: Mitcham, Jennifer
APPLICANT: Skeiky, Yasir
APPLICANT: Persing, David
APPLICANT: Bhatia, Ajay
APPLICANT: Maisonneuve, Jean Francois
APPLICANT: Zhang, Yanni
APPLICANT: Wang, Siqing
APPLICANT: Jen, Shyian
APPLICANT: Lodes, Michael
APPLICANT: Benson, Darin
APPLICANT: Jones, Robert
APPLICANT: Carter, Darlick
APPLICANT: Barth, Brenda
APPLICANT: Douglass, John
TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acnes
FILE REFERENCE: 210121.514C1
CURRENT APPLICATION NUMBER: US/09/978,825

CURRENT FILING DATE: 2003-01-29
NUMBER OF SEQ ID NOS: 30992
SEQ ID NO 12317
LENGTH: 125
TYPE: PRT
ORGANISM: Propionibacterium acnes
US-09-978-825-12317

Query Match 70.9%; Score 436; DB 5; Length 125;
Best Local Similarity 73.8%; Pred. No. 1.6e-29;
Matches 90; Conservative 9; Mismatches 21; Indels 2; Gaps 2;

QY 1 MKLKVYNGTAYDVVDVYKSHENPMGTILFGGCTGAPAPAAAGAGAGAGGEGEIPAPL 60
DB 6 MKLKVYNDVAYDVVDVDTANAPAPAPILFGGAGG-PMKASGG-GAGKAGEGEVAPL 63
QY 61 AGTYSKILVKEGDTVRAGOTVIVLEAMKMETEINAPTDGKVEKYLKERDAVGGGGLIK 120
DB 64 AGTAAKILVABGDVAVKAGVLTLEAMKMETEINAPADGTVKGIIVAVGDVAVGGGGLVA 123
QY 121 IG 122
DB 124 LG 125

RESULT 3
US-10-057-498-12317
Sequence 12317, Application US/10057498
GENERAL INFORMATION:
APPLICANT: Mitcham, Jennifer
APPLICANT: Skeiky, Yasir
APPLICANT: Persing, David
APPLICANT: Lodes, Michael
APPLICANT: Benson, Darin
APPLICANT: Jones, Robert
APPLICANT: Carter, Darlick
APPLICANT: Barth, Brenda
APPLICANT: Douglass, John
TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acne
FILE REFERENCE: 210121.514
CURRENT APPLICATION NUMBER: US/10/057,498
CURRENT FILING DATE: 2001-04-20
NUMBER OF SEQ ID NOS: 29212
SEQ ID NO 12317
LENGTH: 125
TYPE: PRT
ORGANISM: Propionibacterium acnes
US-10-057-498-12317

Query Match 70.9%; Score 436; DB 6; Length 125;
Best Local Similarity 73.8%; Pred. No. 1.6e-29;
Matches 90; Conservative 9; Mismatches 21; Indels 2; Gaps 2;

QY 1 MKLKVYNGTAYDVVDVYKSHENPMGTILFGGCTGAPAPAAAGAGAGAGGEGEIPAPL 60
DB 6 MKLKVYNDVAYDVVDVDTANAPAPAPILFGGAGG-PMKASGG-GAGKAGEGEVAPL 63
QY 61 AGTYSKILVKEGDTVRAGOTVIVLEAMKMETEINAPTDGKVEKYLKERDAVGGGGLIK 120
DB 64 AGTAAKILVABGDVAVKAGVLTLEAMKMETEINAPADGTVKGIIVAVGDVAVGGGGLVA 123
QY 121 IG 122
DB 124 LG 125

RESULT 4
US-10-369-493-10079
Sequence 10079, Application US/10369493
GENERAL INFORMATION:
APPLICANT: Cao, Yongwei
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Goldman, Barry S.
APPLICANT: Chen, Xianfeng
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
FILE REFERENCE: 38-10(52052)B
CURRENT APPLICATION NUMBER: US/10/369,493


```

0y 5 VVWNGTAYDVDDVDKSHENPAGTILFEGGCTG-----APAPAAAGAGACKAGEG-EIP 57
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 481 VEVEGKAFVVKV-----SDGGVSQLTAAAPASAPAAAAPAGAGTPVT 523

```

Qy 58 APLAGTGVKILVKEGDTYKAGOTVLYLEAMKMETEINAPTDGKVEKVLKERDAVGGG 117
Db 524 APLAGTGVKILVKEGDTYKAGOTVLYLEAMKMETEINAPTDGKVEKVLKERDAVGGT 583
Qy 118 LIKI 121
Db 584 LMTL 587

RESULT 7

US-10-369-493-112
Sequence 112, Application US/10369493
GENERAL INFORMATION:
APPLICANT: Cao, Yongwei
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Goldman, Barry S.
APPLICANT: Chen, Xianfeng
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
FILE REFERENCE: 38-10(52052)B
CURRENT APPLICATION NUMBER: US/10/369,493
PRIOR FILING DATE: 2003-02-28
PRIOR APPLICATION NUMBER: US 60/360,039
PRIOR FILING DATE: 2002-02-21
NUMBER OF SEQ ID NOS: 47374
SEQ ID NO 112
LENGTH: 655
TYPE: PRT
ORGANISM: Aquifex aeolicus
US-10-369-493-112

Query Match

Best Local Similarity 31.0%; Score 190.5; DB 6; Length 655;
Matches 49; Conservative 19; Mismatches 52; Indels 13; Gaps 3;

Qy 2 KLKTVNGTAYDV-----VDVDSHE-----NPMGTILFGGGTGGAPAPAGAGAG 49
Db 513 KFKVIRGVSVDASGRPKRYVYINNRLEIQKPKREAIPOGGGAGOTVOSAEERGIP 572
Qy 50 KAGE-GEIPAPLAGTGVKILVKEGDTYKAGOTVLYLEAMKMETEINAPTDGKVEKVLKE 108
Db 573 KATEPGDVTPEPMGKVKIIVKEGEPVQGGTATVTEAMKMEVNAHPIGGIVKIRAKP 632
Qy 109 RDAVGGGGLIKI 121
Db 533 GDVNPDAIMRI 645

RESULT 8

US-10-282-122A-54468
Sequence 54468, Application US/10282122A
GENERAL INFORMATION:
APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-05-23

PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining prior Application data removed - See file wrapper or PATM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 54468
LENGTH: 599
TYPE: PRT
ORGANISM: Campylobacter jejuni
US-10-282-122A-54468

Query Match 30.5%; Score 187.5; DB 6; Length 599;
Best Local Similarity 37.5%; Pred. No. 8.4e-08;
Matches 48; Conservative 15; Mismatches 40; Indels 25; Gaps 3;

Qy 2 KLKTVNGTAYDV-----VDVDSHE-----NPMGTILFGGGTGGAPAPAGAGAG 50
Db 481 KFTYVANGKTYHVEVSGFEPKDVNVSFKVEERK--NIISNSTSSVDA----- 528
Qy 51 AGESEIPAPLAGTGVKILVKEGDTYKAGOTVLYLEAMKMETEINAPTDGKVEKVLKERD 110
Db 529 --ENEVLAGISGNVFKIYVEGEVSGAIVMLEAMKMEIENARKDGIILELCIKIGD 586
Qy 111 AVGGGGL 118
Db 587 TVNDEGLV 594

RESULT 9

US-10-282-122A-77079
Sequence 77079, Application US/10282122A
GENERAL INFORMATION:
APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578

```

; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 77079
; LENGTH: 597
; TYPE: PRT
; ORGANISM: Vibrio cholerae
US-10-282-122A-77079

Query Match          30.1%; Score 185; DB 6; Length 597;
Best Local Similarity 39.3%; Pred. No. 1.4e-07;
Matches 46; Conservative 18; Mismatches 45; Indels 8; Gaps 2;

OY 5 VTWNGTAYDVVDV-DKSHENPMGTILFGGCTGAPAPAGAGAGAGGEGEIPAPLAGTV 64
DB 488 VKVGVGVYDVEV-----GSGGQLTSVYPAGOKAPLAVATPTQGAEA-VAAPLAGTI 539

OY 65 SKILVEGDTYKAGOTVLTLEAMKMETEINAPTDGKVEKVLKERDAYOGGGLIKI 121
DB 540 FKIQEGDVEAGDVLVLLEAMKMETEIRARSGVIELVHKEDSVRYVGASLSTL 596

RESULT 10
US-10-369-493-1263
; Sequence 1263, Application US/10369493
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10152052JB
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 1263
; LENGTH: 571
; TYPE: PRT
; ORGANISM: Pyrococcus horikoshii
US-10-369-493-1263

Query Match          29.9%; Score 184; DB 6; Length 571;
Best Local Similarity 37.5%; Pred. No. 1.6e-07;
Matches 45; Conservative 24; Mismatches 43; Indels 8; Gaps 3;

OY 3 LKVTNGTAYDVVDV-VKSHENPMGTILFGGCTGAPAPAGAGAGAGGEGEIPAPLA 61
DB 458 IKIYNGKEFEYFVGCIEFEPPKPKPY---QAIPOPKREVAPSSV-----VSAPMP 510

OY 62 GTVSKILVEGDTYKAGOTVLTLEAMKMETEINAPTDGKVEKVLKERDAYOGGGLIKI 121
DB 511 GKVLVLVRYVGDVRYVGGLVLLEAMKMETEIRPRDGVVKRIILVKEGAVDTGGLIEL 570

RESULT 11
US-10-369-493-20329
; Sequence 20329, Application US/10369493
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10152052JB
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 20329
; LENGTH: 571
; TYPE: PRT
; ORGANISM: Pyrococcus horikoshii
US-10-369-493-20329

Query Match          29.9%; Score 184; DB 6; Length 571;
Best Local Similarity 37.5%; Pred. No. 1.6e-07;
Matches 45; Conservative 24; Mismatches 43; Indels 8; Gaps 3;

OY 3 LKVTNGTAYDVVDV-VKSHENPMGTILFGGCTGAPAPAGAGAGAGGEGEIPAPLA 61
DB 458 IKIYNGKEFEYFVGCIEFEPPKPKPY---QAIPOPKREVAPSSV-----VSAPMP 510

OY 62 GTVSKILVEGDTYKAGOTVLTLEAMKMETEINAPTDGKVEKVLKERDAYOGGGLIKI 121
DB 511 GKVLVLVRYVGDVRYVGGLVLLEAMKMETEIRPRDGVVKRIILVKEGAVDTGGLIEL 570

RESULT 12
US-10-369-493-13831
; Sequence 13831, Application US/10369493
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10152052JB
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 13831
; LENGTH: 1076
; TYPE: PRT
; ORGANISM: Pseudomonas fluorescens
US-10-369-493-13831

Query Match          29.3%; Score 180.5; DB 6; Length 1076;
Best Local Similarity 34.6%; Pred. No. 6.6e-07;
Matches 47; Conservative 17; Mismatches 43; Indels 29; Gaps 3;

OY 5 VTWNGTAYDVVDV-----DVDRKSHENPMGTILFGGCTGAPAPAGAG 45
DB 950 IDVHGERTYRVDITGVCAEGRHRYLTSDGNPEVEVEPLNEFVGSSSKRKAASA--- 1006

OY 46 AGAGRAGEGEIPAPLAGTGVSKILVEGDTYKAGOTVLTLEAMKMETEINAPTDGKVEKVL 105
DB 1007 -----PGHVSITPMGNIVDLVKEGDTYKAGAVLTLEAMKMETEVOAAIAGKVTATIH 1059

OY 106 VKERDAYOGGGLIKI 121
DB 1060 VAKGDRVNPGETLIEI 1075

RESULT 13
US-10-369-493-14349
; Sequence 14349, Application US/10369493
```

```

; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 14349
; LENGTH: 574
; TYPE: PRT
; ORGANISM: Agrobacterium tumefaciens
US-10-369-493-14349

Query Match      28.7%; Score 176.5; DB 6; Length 574;
Best Local Similarity 43.8%; Pred. No. 6.7e-07;
Matches 42; Conservative 17; Mismatches 28; Indels 9; Gaps 3;
```

```

; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 15017
; LENGTH: 574
; TYPE: PRT
; ORGANISM: Agrobacterium tumefaciens
US-10-369-493-15017

Query Match      28.7%; Score 176.5; DB 6; Length 574;
Best Local Similarity 43.8%; Pred. No. 6.7e-07;
Matches 42; Conservative 17; Mismatches 28; Indels 9; Gaps 3;
```

```

; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 15017
; LENGTH: 574
; TYPE: PRT
; ORGANISM: Agrobacterium tumefaciens
US-10-369-493-15017

Query Match      28.7%; Score 176.5; DB 6; Length 574;
Best Local Similarity 43.8%; Pred. No. 6.7e-07;
Matches 42; Conservative 17; Mismatches 28; Indels 9; Gaps 3;
```

```

; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 15017
; LENGTH: 574
; TYPE: PRT
; ORGANISM: Agrobacterium tumefaciens
US-10-369-493-15017

Query Match      28.7%; Score 176.5; DB 6; Length 574;
Best Local Similarity 43.8%; Pred. No. 6.7e-07;
Matches 42; Conservative 17; Mismatches 28; Indels 9; Gaps 3;
```

```

; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 11591
; LENGTH: 575
; TYPE: PRT
; ORGANISM: Agrobacterium tumefaciens
US-10-369-493-11591

Query Match      28.7%; Score 176.5; DB 6; Length 575;
Best Local Similarity 43.8%; Pred. No. 6.8e-07;
Matches 42; Conservative 17; Mismatches 28; Indels 9; Gaps 3;
```

```

; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 14608
; LENGTH: 575
; TYPE: PRT
; ORGANISM: Agrobacterium tumefaciens
US-10-369-493-14608

Query Match      28.7%; Score 176.5; DB 6; Length 575;
Best Local Similarity 43.8%; Pred. No. 6.8e-07;
Matches 42; Conservative 17; Mismatches 28; Indels 9; Gaps 3;
```

```

; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 14608
; LENGTH: 575
; TYPE: PRT
; ORGANISM: Agrobacterium tumefaciens
US-10-369-493-14608

Query Match      28.7%; Score 176.5; DB 6; Length 575;
Best Local Similarity 43.8%; Pred. No. 6.8e-07;
Matches 42; Conservative 17; Mismatches 28; Indels 9; Gaps 3;
```

```

; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
US-09-134-000C-6221
; Sequence 6221, Application US/09134000C
; GENERAL INFORMATION:
```

```

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; FILE REFERENCE: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09/134,000C
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6221
; LENGTH: 147
; TYPE: PRN
; ORGANISM: Enterococcus faecalis
US-09-134-000C-6221

Query Match          28.5%; Score 175.5; DB 5; Length 147;
Best Local Similarity 37.2%; Pred. No. 1.6e-07;
Matches 48; Conservative 17; Mismatches 55; Indels 9; Gaps 2;

QY 2 KLVYNGTAYDVVD-----VDKSHENPMGTILFGGTCGAPAPAGAGAGAGAG 52
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 18 KFKISIDGKEVLMEMIGVQPPAPVAPQPTAVATTETPPAPAVEETPPASAPAPAG 77
    ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 53 EGEIPAPLACTVSKILYKEDDTYKAGTVLVLEAMKMETEINAPTDGKVEKVLKRDV 112
    ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 78 ADAPAPMPCTVTKVLVNVGDTVSENOPLILLEMKMETEIVAKGAGTVGIHTQGIIV 137
    ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY 113 QGGGGLIKI 121
    |:::|
Db 138 NPGEPILITI 146

RESULT 18
; Sequence 6221, Application US/09134000C
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; FILE REFERENCE: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09/134,000C
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6221
; LENGTH: 147
; TYPE: PRN
; ORGANISM: Enterococcus faecalis
US-09-134-000C-6221

Query Match          28.5%; Score 175.5; DB 5; Length 147;
Best Local Similarity 37.2%; Pred. No. 1.6e-07;
Matches 48; Conservative 17; Mismatches 55; Indels 9; Gaps 2;

QY 2 KLVYNGTAYDVVD-----VDKSHENPMGTILFGGTCGAPAPAGAGAGAG 52
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 18 KFKISIDGKEVLMEMIGVQPPAPVAPQPTAVATTETPPAPAVEETPPASAPAPAG 77
    ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 53 EGEIPAPLACTVSKILYKEDDTYKAGTVLVLEAMKMETEINAPTDGKVEKVLKRDV 112
    ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 78 ADAPAPMPCTVTKVLVNVGDTVSENOPLILLEMKMETEIVAKGAGTVGIHTQGIIV 137
    ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY 113 QGGGGLIKI 121
    |:::|
Db 138 NPGEPILITI 146

RESULT 19
US-10-282-122A-67750
; Sequence 67750, Application US/10282122A
; GENERAL INFORMATION:
```

```

; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA, 034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; PRIOR FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining prior Application data removed - See file Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 67750
; LENGTH: 602
; TYPE: PRN
; ORGANISM: Pseudomonas putida
US-10-282-122A-67750

Query Match          28.5%; Score 175.5; DB 6; Length 602;
Best Local Similarity 33.8%; Pred. No. 8.7e-07;
Matches 46; Conservative 17; Mismatches 44; Indels 29; Gaps 3;

QY 5 VTVNGTAYDVVD-----VDKSHENPMGTILFGGTCGAPAPAGAG 45
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 475 IDVNGEYTRVDITGVGKAEKRRHFTYLSIDGMPREVEYFEPINFEVSGGSKRRA----- 529
    ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 46 AGACKAGEGEIPAPLACTVSKILYKEDDTYKAGTVLVLEAMKMETEINAPTDGKVEKVL 105
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 530 -----TDPGHVSFTMPGNITVDVLVKEGDYKAGCAVGLITEAMKMETEVOAIAKGVAAIH 584
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY 106 VKERDAVQGGGLIKI 121
    |:::|
Db 585 VAKGDRVTPGEILTEI 600

RESULT 20
US-10-369-493-18634
; Sequence 18634, Application US/10369493
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
```

```
FILE REFERENCE: 38-10(52052)B
CURRENT APPLICATION NUMBER: US/10/369,493
CURRENT FILING DATE: 2003-02-28
PRIOR APPLICATION NUMBER: US 60/360,039
PRIOR FILING DATE: 2002-02-21
NUMBER OF SEQ ID NOS: 47374
SEQ ID NO 18634
LENGTH: 610
TYPE: PRT
ORGANISM: Halobacterium sp. NRC-1
US-10-369-493-18634

Query Match
Best Local Similarity 38.1%; Pred. No. 8.8e-07; Length 610;
Matches 48; Conservative 18; Mismatches 43; Indels 17; Gaps 5;

QY 5 VTWNGTAYDVVDVKNKSHENPMGTILFG-----GTGAPAPAPAGAGAGAKA---GEGE 55
Db 492 VEVNGKREFEYNLE-----ERGAAGFAPEADTGGGGPPEP-AGGADDETVVEGSGET 543

QY 56 IPAPLAVTSKILYKEDDTYKAGQTVLYLEAMKMETINAPTDGKVEKVLVKERDAVOGG 115
Db 544 VTAEMOGTIIDVAVSEGDVADAGDVLVLEAMKMENTVVAHSGGTVTVQVAVSEDDSYDMD 603

QY 116 QGLIKI 121
Db 604 DVLVVI 609

RESULT 21
US-10-282-122A-61290
Sequence 61290, Application US/10282122A
GENERAL INFORMATION:
APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See file Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 61290
```

```
LENGTH: 596
TYPE: PRT
ORGANISM: Legionella pneumophila
US-10-282-122A-61290

Query Match
Best Local Similarity 28.5%; Score 175; DB 6; Length 596;
Matches 37; Conservative 12; Mismatches 23; Indels 0; Gaps 0;

QY 50 KAGEEIPAPLAVTSKILYKEDDTYKAGQTVLYLEAMKMETINAPTDGKVEKVLVKER 109
Db 522 KIGSGDITVAIPSSIIIAIHVSDEYKAGQAVLYLEAMKMETIKAPANGVAEILICQNG 581

QY 110 DAVOGGGLIKI 121
Db 582 DKVTGQVLIRV 593

RESULT 22
US-10-282-122A-46866
Sequence 46866, Application US/10282122A
GENERAL INFORMATION:
APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See file Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 46866
LENGTH: 1148
TYPE: PRT
ORGANISM: Bacillus anthracis
US-10-282-122A-46866

Query Match
Best Local Similarity 28.5%; Score 175; DB 6; Length 1148;
Matches 37; Conservative 14; Mismatches 31; Indels 0; Gaps 0;

QY 40 APAGAGAGKAGCEGEPAPLAVTSKILYKEDDTYKAGQTVLYLEAMKMETINAPTDG 99
```

Db 1065 ATVAGRVGNENPHNHSATMGVYIKVYKESGDEVKKGDSMTAEAKMETTVQAPENG 1124
QY 100 KVEKLVKENDAVOGGGLIKI 121
Db 1125 KVKKYVNDGDAICGTGDLIEL 1146

RESULT 23
US-10-089-057-4
Sequence 4, Application US/10089057
GENERAL INFORMATION:
APPLICANT: HIRANO, SEIKO
APPLICANT: NONAKA, GEN
APPLICANT: MATSUZAKI, YUMI
APPLICANT: AKIYOSHI, NAOKI
APPLICANT: NAKAMURA, KANAE
APPLICANT: KIMURA, EICHIRO
APPLICANT: ODOMI, TSUYOSHI
APPLICANT: MATSUI, KAZUHIKO
APPLICANT: KAWAHARA, YOSHIO
APPLICANT: KURAHASHI, OSAMU
APPLICANT: NAKAMATSU, TSUYOSHI
APPLICANT: SUDIMOTO, SHINICHI
TITLE OF INVENTION: GENES FOR HEAT-RESISTANT ENZYMES OF AMINO ACID BIOSYNTHETIC PATH
TITLE OF INVENTION: DERIVED FROM THERMOPHILIC CORYNEFORM BACTERIA
FILE REFERENCE: 221519USOPT
CURRENT FILING DATE: 2002-04-03
CURRENT FILING DATE: 2002-04-03
PRIOR FILING DATE: 2000-10-04
PRIOR FILING DATE: 2000-10-04
PRIOR FILING DATE: 2000-10-04
PRIOR FILING DATE: 1999-10-04
PRIOR FILING DATE: 1999-11-01
PRIOR FILING DATE: 1999-11-01
PRIOR FILING DATE: 2000-04-21
NUMBER OF SEQ ID NOS: 129
SOFTWARE: PatentIn version 3.1
SEQ ID NO 4
LENGTH: 591
TYPE: PRT
ORGANISM: Corynebacterium thermocaulnogenes
US-10-089-057-4

Query Match 28.3%; Score 174; DB 6; Length 591;
Best Local Similarity 36.3%; Pred. No. 1.1e-06;
Matches 45; Conservative 20; Mismatches 45; Indels 14; Gaps 3;

QY 2 KIKTVNGTAVDVVDVSKSHENPMGTLIFGGGTGAPAP-----AAGGAGAGAGGEGTIP 57
Db 477 KVIYEIDRRVVALP-----GDALGGGAGAAKKPKRRRAGGAKAGVSGD-SVA 526
QY 58 APLATGVSILIKEDGTVKAGOTVLVLEAMKMETEINAPTDGKVEKVLKENDAVOGGOG 117
Db 527 APMGCTIVKVNEDGAEVSEGTIVVLEAMKMETEINAPTDGKVEKVLKENDAVOGGOG 117

QY 118 LIRI 121
Db 587 LLEI 590

RESULT 24
US-10-169-048-16
Sequence 16, Application US/10169048
GENERAL INFORMATION:
APPLICANT: Clarke, Edna Elizabeth
APPLICANT: Zhou, Liqing
APPLICANT: Shea, Jacqueline Elizabeth
APPLICANT: Feldman, Robert Graham
APPLICANT: Holden, David William
TITLE OF INVENTION: Streptococcus Pyogenes Virulence Genes and Proteins And Their Use
FILE REFERENCE: GIE-97
CURRENT APPLICATION NUMBER: US/10/169,048

;; CURRENT FILING DATE: 2002-06-24
;; PRIOR APPLICATION NUMBER: PCT/GB00/04997
;; PRIOR FILING DATE: 2000-12-22
;; NUMBER OF SEQ ID NOS: 62
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO 16
;; LENGTH: 134
;; TYPE: PRT
;; ORGANISM: Streptococcus pyogenes
US-10-169-048-16

Query Match 28.2%; Score 173.5; DB 6; Length 134;
Best Local Similarity 32.4%; Pred. No. 2.1e-07;
Matches 47; Conservative 15; Mismatches 42; Indels 41; Gaps 2;

QY 2 KIKTVNGTAVDVVDVSKSHENPMGTLIFGGGTGAPAPAGCA----- 46
Db 5 KFKITIDGKEYLVE-----ELGAPQAAAPAPPISTPVPTPEAS 47
QY 47 -----GAGKAGEGEIPAPLACTVSKILVKEGDTVKAGOTVLVLEAMKMETEINAPT 97
Db 48 QVEAQAQPVAAAGADAPSPMPTILKVLAVGDQYTENQPLILEAMKMETEINAVAS 107

QY 98 DKEKVLKENDAVOGGGLIKIG 122
Db 108 AGTITAIHVPGQVYVNPBGGLITIG 132

RESULT 25
US-10-282-122A-75568
Sequence 75568, Application US/10282122A
GENERAL INFORMATION:
APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR FILING DATE: 2000-03-21
PRIOR FILING DATE: 2000-03-21
PRIOR FILING DATE: 2000-05-23
PRIOR FILING DATE: 2000-05-26
PRIOR FILING DATE: 2000-05-26
PRIOR FILING DATE: 2000-09-06
PRIOR FILING DATE: 2000-09-06
PRIOR FILING DATE: 2000-09-09
PRIOR FILING DATE: 2000-12-22
PRIOR FILING DATE: 2000-12-22
PRIOR FILING DATE: 2001-02-09
PRIOR FILING DATE: 2001-02-09
PRIOR FILING DATE: 2001-02-16
Remainding Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 75568
LENGTH: 591

```

; TYPE: PRT
; ORGANISM: Salmonella typhi
US-10-282-122A-75368
Query Match
Best Local Similarity 40.2%; Score 173.5; DB 6; Length 591;
Matches 47; Conservative 14; Mismatches 49; Indels 7; Gaps 3;

OY 5 VTWNGTAVDVVDVDSKHEHNMGTILFGGCTGAPAPAGGAGAGKAGEEIPAPLAGTV 64
DB 481 VEEGKAFVVRVS--DGGDISQLTTAV--PAASSAPVOAAAPAGAGT-----PVTAPLAGNI 533

OY 65 SKIIVKEDGVKAGOTVLEAMKMETEINAPTDGKVEKYLVERDAVGGGGLIKI 121
DB 534 WKVIATGOSVABGDVLLLEAMKMETEIRAAQGTVRGIAVKSGDVSVGDTLMTL 590

RESULT 26
US-10-282-122A-76334
; Sequence 76334, Application US/10282122A
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See file wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 76334
; LENGTH: 591
; TYPE: PRT
; ORGANISM: Salmonella typhi
US-10-282-122A-76334

Query Match
Best Local Similarity 40.2%; Score 173.5; DB 6; Length 591;
Matches 47; Conservative 14; Mismatches 49; Indels 7; Gaps 3;

OY 5 VTWNGTAVDVVDVDSKHEHNMGTILFGGCTGAPAPAGGAGAGKAGEEIPAPLAGTV 64
DB 481 VEEGKAFVVRVS--DGGDISQLTTAV--PAASSAPVOAAAPAGAGT-----PVTAPLAGNI 533
```

```

OY 65 SKIIVKEDGVKAGOTVLEAMKMETEINAPTDGKVEKYLVERDAVGGGGLIKI 121
DB 534 WKVIATGOSVABGDVLLLEAMKMETEIRAAQGTVRGIAVKSGDVSVGDTLMTL 590

RESULT 27
US-10-369-493-1053
; Sequence 1053, Application US/10369493
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(53052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 1053
; LENGTH: 567
; TYPE: PRT
; ORGANISM: Methanococcus jannaschii
US-10-369-493-1053

Query Match
Best Local Similarity 50.7%; Score 173; DB 6; Length 567;
Matches 35; Conservative 12; Mismatches 22; Indels 0; Gaps 0;

OY 53 EGEIPAPLAGTVKSKIVKEDGVKAGOTVLEAMKMETEINAPTDGKVEKYLVERDAV 112
DB 498 EGAVTSPFRMVKIKYKEDGKKGVDVYLLAMKMEHIESPVEGYERIIIDEGDAV 557

OY 113 QGGGGLIKI 121
DB 558 NVGDVIMII 566

RESULT 28
US-10-282-122A-69598
; Sequence 69598, Application US/10282122A
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
```



```

PRIORITY APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 52088
LENGTH: 1144
TYPE: PRT
ORGANISM: Clostridium acetobutylicum
US-10-282-122A-52088

Query Match          27.6%  Score 170; DB 6; Length 1144;
Best Local Similarity 47.3%; Pred. No. 5.4e-06;
Matches 35; Conservative 13; Mismatches 26; Indels 0; Gaps 0;

QY      48  AGKAGEGEPPLPLAGTVSKILVKEGDTYVAKAGQTVLYLEPAMKEFEINAPDGKVEKVLVK 107
       1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      1070 ADSSKKRKGISIPGNVYKVVKPPEDKVKKKDSLMLVIMAKMETVWSVEDGTGCTIVK 1129
       1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      108 ERDAVGGGGLIKI 121
       1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      1130 EGDQVSGQLVTKL 1143
       1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 30
US-10-282-122A-43666
; Sequence 43666, Application US/10282122A
; GENERAL INFORMATION:
; APPLICANT: Wang, Liansu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITR.034A
; CURRENT APPLICATION NUMBER: US/10/282.122A
; PRIOR FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 43666
; LENGTH: 607
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; US-10-282-122A-43666

```

Query Match 27.48; Score 168.5; DB 6; Length 607;
 Best Local Similarity 32.48; Pred. No. 3,4e-06;
 Matches 44; Conservative 20; Mismatches 43; Indels 29; Gaps 3;

```

QY 5 VTNGTAVDVV-----DVDRSHE-----NPMGTILFEGGTGAPAPAAAG 45
   :|:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 480 IDVHGFTYRVDTITGVGVKSDNKRHFYLSIDGMPEEYVFEPLNEVYAGSASGRKHA----- 534
   :|:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 46 AGACKAGEGETPAPLACTVSKILVKEGSDTVKAGQTVLLEAMKMETEINAPTDGKVEKVL 105
   :|:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 535 -----SEPGHVTMPGNITVDVLVKEGDSVAKGQAVLITEAMKMETEYQAGIAGTVKAIH 589
   :|:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 106 VKERDAVQGGGGLIKI 121
   :|:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 590 VAKGDRVNPGEILEI 605
   :|:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
    
```

Search completed: May 1, 2003, 08:09:54
 Job time : 410.844 secs

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 1, 2003, 07:52:32 ; Search time 235.156 Seconds
(Without alignments)
50.445 Million cell updates/sec

Title: US-09-987-485-2

Perfect score: 342

Sequence: 1 EGRIPAPLAGTYSKILVKEG.....KVLVKEPDVAGGGLIKIG 70

Scoring table: BLOSUM62

Searched: 777667 seqs, 169465163 residues

Total number of hits satisfying chosen parameters: 777667

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 100 summaries

Database : Pending_Patents_AA_New.*
1: /cgn2_6/ptodata/2/paa/PCT_NEW_COMB.pep.*
2: /cgn2_6/ptodata/2/paa/US06_NEW_COMB.pep.*
3: /cgn2_6/ptodata/2/paa/US07_NEW_COMB.pep.*
4: /cgn2_6/ptodata/2/paa/US08_NEW_COMB.pep.*
5: /cgn2_6/ptodata/2/paa/US09_NEW_COMB.pep.*
6: /cgn2_6/ptodata/2/paa/US10_NEW_COMB.pep.*
7: /cgn2_6/ptodata/2/paa/US60_NEW_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match Length	DB ID	Description
1	263	76.9	125	1 PCT-US02-33727-12317	Sequence 12317, A
2	263	76.9	125	5 US-09-978-825-12317	Sequence 12317, A
3	263	76.9	125	6 US-10-057-498-12317	Sequence 12317, A
4	180	52.6	571	6 US-10-369-493-1263	Sequence 1263, AP
5	180	52.6	571	6 US-10-369-493-20329	Sequence 20329, A
6	174	50.9	655	6 US-10-369-493-112	Sequence 112, App
7	174	50.9	1148	6 US-10-282-122A-46866	Sequence 46866, A
8	173	50.6	567	6 US-10-369-493-1053	Sequence 1053, AP
9	172	50.3	143	6 US-10-282-122A-48959	Sequence 48959, A
10	167	48.8	1076	6 US-10-369-493-13831	Sequence 13831, A
11	167	48.8	1144	6 US-10-282-122A-52088	Sequence 52088, A
12	166	48.5	586	6 US-10-282-122A-61290	Sequence 61290, A
13	165	48.2	598	6 US-10-282-122A-60140	Sequence 60140, A
14	164	48.0	597	6 US-10-282-122A-77079	Sequence 77079, A
15	163	47.9	599	6 US-10-282-122A-54468	Sequence 54468, A
16	160.5	46.9	75	6 US-10-224-539A-7	Sequence 7, App1
17	160.5	46.9	75	6 US-10-370-959-133	Sequence 133, App
18	160.5	46.9	75	6 US-10-377-097-133	Sequence 133, App
19	160	46.8	502	6 US-10-282-122A-69598	Sequence 69598, A
20	159.5	46.6	574	6 US-10-369-493-14349	Sequence 14349, A
21	159.5	46.6	574	6 US-10-369-493-15017	Sequence 15017, A
22	159.5	46.6	575	6 US-10-369-493-11591	Sequence 11591, A
23	159.5	46.6	575	6 US-10-369-493-14608	Sequence 14608, A
24	159	46.5	602	6 US-10-282-122A-67750	Sequence 67750, A
25	158	46.2	607	6 US-10-282-122A-43666	Sequence 43666, A
26	158	46.2	612	6 US-10-366-683-19134	Sequence 19134, A

27	158	46.2	612	6 US-10-419-128-19134	Sequence 19134, A
28	157	45.9	1146	6 US-10-282-122A-60488	Sequence 60488, A
29	157	45.9	1150	6 US-10-369-493-17338	Sequence 17338, A
30	155.5	45.5	1141	6 US-10-282-122A-53885	Sequence 53885, A
31	153	44.7	620	6 US-10-369-493-119	Sequence 119, App
32	153	44.7	1144	6 US-10-282-122A-52777	Sequence 52777, A
33	151.5	44.3	589	6 US-10-369-493-10079	Sequence 10079, A
34	151	44.2	151	6 US-10-282-122A-74706	Sequence 74706, A
35	151	44.2	1142	6 US-10-282-122A-57942	Sequence 57942, A
36	150	43.9	190	6 US-10-417-884-5720	Sequence 5720, AP
37	150	43.9	1143	6 US-10-282-122A-52917	Sequence 52917, A
38	148	43.3	162	6 US-10-282-122A-72184	Sequence 72184, A
39	148	43.3	591	6 US-10-282-122A-75568	Sequence 75568, A
40	148	43.3	591	6 US-10-282-122A-76334	Sequence 76334, A
41	148	43.3	1147	6 US-10-369-493-11450	Sequence 11450, A
42	148	43.3	1151	6 US-10-369-493-14817	Sequence 14817, A
43	148	43.3	1151	6 US-10-369-493-15006	Sequence 15006, A
44	147	43.0	134	6 US-10-169-448-16	Sequence 16, App1
45	146	42.7	611	6 US-10-282-122A-70190	Sequence 70190, A
46	146	42.7	989	5 US-09-950-084-3973	Sequence 3973, AP
47	146	42.7	1150	6 US-10-282-122A-44391	Sequence 44391, A
48	145	42.4	701	6 US-10-366-683-27999	Sequence 27999, A
49	145	42.4	701	6 US-10-419-128-27999	Sequence 27999, A
50	144	42.1	108	5 US-09-880-505A-166	Sequence 166, App
51	144	42.1	147	5 US-09-134-000C-6221	Sequence 6221, AP
52	144	42.1	147	5 US-09-134-000C-6221	Sequence 6221, AP
53	144	42.1	161	6 US-10-282-122A-73808	Sequence 73808, A
54	144	42.1	654	6 US-10-369-493-15564	Sequence 15564, A
55	144	42.1	654	6 US-10-369-493-15937	Sequence 15937, A
56	144	42.1	1142	6 US-10-282-122A-42528	Sequence 42528, A
57	144	42.1	1163	5 US-09-134-000C-5707	Sequence 5707, AP
58	144	42.1	1163	5 US-09-134-000C-5707	Sequence 5707, AP
59	143	41.8	725	6 US-10-224-539A-2	Sequence 2, App1
60	143	41.8	725	6 US-10-224-539A-9	Sequence 9, App1
61	143	41.8	725	6 US-10-370-959-128	Sequence 128, App
62	143	41.8	725	6 US-10-370-959-135	Sequence 135, App
63	143	41.8	725	6 US-10-377-097-128	Sequence 128, App
64	143	41.8	725	6 US-10-377-097-135	Sequence 135, App
65	143	41.8	725	6 US-60-453-135-10473	Sequence 10473, A
66	143	41.8	725	6 US-60-453-135-10473	Sequence 10473, A
67	143	41.8	1137	6 US-10-369-493-18351	Sequence 18351, A
68	143	41.8	1167	6 US-10-369-493-22819	Sequence 22819, A
69	143	41.8	1185	6 US-10-369-493-2488	Sequence 2488, AP
70	142	41.5	602	6 US-10-282-122A-67292	Sequence 67292, A
71	142	41.5	901	6 US-10-369-493-17056	Sequence 17056, A
72	141	41.2	1078	6 US-10-369-493-17039	Sequence 17039, A
73	140	40.9	568	6 US-10-369-493-1203	Sequence 1203, AP
74	140	40.9	1152	6 US-10-369-493-12027	Sequence 12027, A
75	139.5	40.8	162	6 US-10-282-122A-42491	Sequence 42491, A
76	139.5	40.8	168	5 US-09-134-000C-4552	Sequence 4552, AP
77	139.5	40.8	168	5 US-09-134-000C-4552	Sequence 4552, AP
78	139.5	40.8	1127	6 US-10-282-122A-62684	Sequence 62684, A
79	139.5	40.8	1127	6 US-10-282-122A-64810	Sequence 64810, A
80	138	40.4	611	6 US-10-282-122A-70702	Sequence 70702, A
81	138	40.4	1184	6 US-10-092-411A-3478	Sequence 3478, AP
82	138	40.4	1184	6 US-10-369-493-1491	Sequence 1491, AP
83	137	40.1	650	6 US-10-369-493-13692	Sequence 13692, A
84	136.5	39.9	158	6 US-10-282-122A-52530	Sequence 52530, A
85	135	39.5	640	6 US-10-366-683-213696	Sequence 213696, A
86	135	39.5	676	6 US-10-366-683-26143	Sequence 26143, A
87	135	39.5	676	6 US-10-419-128-26143	Sequence 26143, A
88	135	39.5	1797	6 US-10-369-493-127	Sequence 127, App
89	134	39.2	1144	1 PCT-US02-36123-2136	Sequence 2136, App
90	133.5	39.0	159	6 US-10-282-122A-51940	Sequence 51940, A
91	133	38.9	611	6 US-10-282-122A-71785	Sequence 71785, A
92	133	38.9	660	6 US-10-369-493-8253	Sequence 8253, AP
93	133	38.9	1178	6 US-10-369-493-18199	Sequence 18199, A
94	133	38.9	1194	6 US-10-369-493-17337	Sequence 17337, A
95	132	38.6	588	6 US-10-282-122A-73367	Sequence 73367, A
96	131.5	38.5	156	6 US-10-282-122A-49663	Sequence 49663, A
97	131	38.3	716	6 US-10-369-493-12942	Sequence 12942, A
98	130.5	38.2	91	6 US-10-282-122A-47349	Sequence 47349, A
99	130.5	38.2	654	6 US-10-369-493-16934	Sequence 16934, A

100 130 38.0 598 6 US-10-282-122A-63744

Sequence 63744, A

ALIGNMENTS

RESULT 1

PCT-US02-32727-12317
; Sequence 12317, Application PC/TUS0232727
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer
; APPLICANT: Skelky, Yasir
; APPLICANT: Persing, David
; APPLICANT: Bhalla, Ajay
; APPLICANT: Maisonneuve, Jean Francois
; APPLICANT: Zhang, Yanni
; APPLICANT: Wang, Siqing
; APPLICANT: Jen, Shylian
; APPLICANT: Lodes, Michael
; APPLICANT: Benson, Darin
; APPLICANT: Jones, Robert
; APPLICANT: Carter, Darlick
; APPLICANT: Barth, Brenda
; APPLICANT: Douglass, John
; TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acnes
; FILE REFERENCE: 210121.514C1
; CURRENT APPLICATION NUMBER: PCT/US02/32727
; CURRENT FILING DATE: 2002-10-11
; NUMBER OF SEQ ID NOS: 30992
; SEQ ID NO 12317
; LENGTH: 125
; TYPE: PRT
; ORGANISM: Propionl acnes
PCT-US02-32727-12317

Query Match 76.9%; Score 263; DB 1; Length 125;
Best Local Similarity 74.3%; Pred. No. 5.2e-23;

Matches 52; Conservative 7; Mismatches 11; Indels 0; Gaps 0;

QY 1 EGEIPAPLAGTVSKILVEGDTVKAGQTVLVLEAKMKMETEINAPTDGKVEKVLVERDAY 60

DB 56 EGEVAPPLAGTVAKTILVAGDAVKAGQVLTLEAKMKMETEINAPDGTGKILVAVGDV 115

QY 61 OGCGGLIKTG 70

DB 116 OGCGGLVALG 125

RESULT 2

US-09-978-825-12317
; Sequence 12317, Application US/09978825
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer
; APPLICANT: Skelky, Yasir
; APPLICANT: Persing, David
; APPLICANT: Bhalla, Ajay
; APPLICANT: Maisonneuve, Jean Francois
; APPLICANT: Zhang, Yanni
; APPLICANT: Wang, Siqing
; APPLICANT: Jen, Shylian
; APPLICANT: Lodes, Michael
; APPLICANT: Benson, Darin
; APPLICANT: Jones, Robert
; APPLICANT: Carter, Darlick
; APPLICANT: Barth, Brenda
; APPLICANT: Douglass, John
; TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acnes
; FILE REFERENCE: 210121.514C1
; CURRENT APPLICATION NUMBER: US/09/978, 825
; CURRENT FILING DATE: 2003-01-29
; NUMBER OF SEQ ID NOS: 30992
; SEQ ID NO 12317
; LENGTH: 125

; TYPE: PRT
; ORGANISM: Propionl acnes
US-09-978-825-12317

Query Match 76.9%; Score 263; DB 5; Length 125;
Best Local Similarity 74.3%; Pred. No. 5.2e-23;

Matches 52; Conservative 7; Mismatches 11; Indels 0; Gaps 0;

QY 1 EGEIPAPLAGTVSKILVEGDTVKAGQTVLVLEAKMKMETEINAPTDGKVEKVLVERDAY 60

DB 56 EGEVAPPLAGTVAKTILVAGDAVKAGQVLTLEAKMKMETEINAPDGTGKILVAVGDV 115

QY 61 OGCGGLIKTG 70

DB 116 OGCGGLVALG 125

RESULT 3

US-10-057-498-12317
; Sequence 12317, Application US/10057498
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer
; APPLICANT: Skelky, Yasir
; APPLICANT: Persing, David
; APPLICANT: Bhalla, Ajay
; APPLICANT: Maisonneuve, Jean Francois
; APPLICANT: Zhang, Yanni
; APPLICANT: Wang, Siqing
; APPLICANT: Jen, Shylian
; APPLICANT: Lodes, Michael
; APPLICANT: Benson, Darin
; APPLICANT: Jones, Robert
; APPLICANT: Carter, Darlick
; APPLICANT: Barth, Brenda
; APPLICANT: Douglass, John
; TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acnes
; FILE REFERENCE: 210121.514
; CURRENT APPLICATION NUMBER: US/10/057, 498
; CURRENT FILING DATE: 2001-04-20
; NUMBER OF SEQ ID NOS: 29212
; SEQ ID NO 12317
; LENGTH: 125
; TYPE: PRT
; ORGANISM: Propionl acnes
US-10-057-498-12317

Query Match 76.9%; Score 263; DB 6; Length 125;
Best Local Similarity 74.3%; Pred. No. 5.2e-23;

Matches 52; Conservative 7; Mismatches 11; Indels 0; Gaps 0;

QY 1 EGEIPAPLAGTVSKILVEGDTVKAGQTVLVLEAKMKMETEINAPTDGKVEKVLVERDAY 60

DB 56 EGEVAPPLAGTVAKTILVAGDAVKAGQVLTLEAKMKMETEINAPDGTGKILVAVGDV 115

QY 61 OGCGGLIKTG 70

DB 116 OGCGGLVALG 125

RESULT 4

US-10-369-493-1263
; Sequence 1263, Application US/10369493
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369, 493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360, 039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 1263
; LENGTH: 571
; TYPE: PRT
; ORGANISM: Pyrococcus horikoshii
US-10-369-493-1263

Query Match 52.6%; Score 180; DB 6; Length 571;
Best Local Similarity 54.5%; Pred. No. 1.8e-12;

Matches	36;	Conservative	14;	Mismatches	16;	Indels	0;	Gaps	0;
QY	4	IPAPLGYTKLTVGEGDVGKGVLYLVLAKKMEETAPDQKREKLYVERADVGG	63						
Db	505	VSAPHPGKVLRLVLYGVGRVYGGGLTVLEAKKMEETLSPDQGVYKRLVLEGEANDYG	564						
QY	64	QGLIKI	69						
Db	565	QPLIEL	570						

RESULT 5
US-10-369-493-20329

```

: Sequence 20329, Application US/10369493
: GENERAL INFORMATION:
: APPLICANT: Cao, Yongwei
: APPLICANT: Hinkle, Gregory J.
: APPLICANT: Slater, Steven C.
: APPLICANT: Goldman, Barry S.
: APPLICANT: Chen, Xianfeng
: TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
: TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
: FILE REFERENCE: 38-10(52052)B
: CURRENT APPLICATION NUMBER: US/10/369,493
: CURRENT FILING DATE: 2003-02-28
: PRIOR APPLICATION NUMBER: US 60/360,039
: PRIOR FILING DATE: 2002-02-21
: NUMBER OF SEQ ID NOS: 47374
: SEQ ID NO 20329
: LENGTH: 571
: TYPE: prt
: ORGANISM: Pyrococcus horikoshii
: US-10-369-493-20329

```

Query Match	52.6%;	Score 180;	DB 6;	Length 571;
Best Local Similarity	54.5%;	Pred. No. 1.8e-12;		
Matches 36;	Conservative 14;	Mismatches 16;	Indels 0;	Gaps 0

RESULT 6
US-10-369-493-112

```

: Sequence 112, Application US/10369493
: GENERAL INFORMATION:
: APPLICANT: Cao, Yongwei
: APPLICANT: Hinkle, Gregory J.
: APPLICANT: Slater, Steven C.
: APPLICANT: Goldman, Barry S.
: APPLICANT: Chen, Xiandeng
: TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
: TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
: FILE REFERENCE: 38-10(52052)B
: CURRENT APPLICATION NUMBER: US/10/369,493
: CURRENT FILING DATE: 2003-02-28
: PRIOR APPLICATION NUMBER: US 60/360,039
: PRIOR FILING DATE: 2002-02-21
: NUMBER OF SEQ ID NOS: 47374
: SEQ ID NO 112
: LENGTH: 655
: TYPE: prt
: ORGANISM: Aquifex aeolicus
US-10-369-493-112

```

Query Match	50.9%;	Score 174;	DB 6;	Length 655;
Best Local Similarity	50.0%;	Pred. No. 1e-11;		
Matches 34;	Conservative 13;	Mismatches 21;	Indels 0;	Gaps 0

[illegible]

RESULT 7
DS-10-28

```

Sequence 46866, Application US/10282122A
GENERAL INFORMATION:
APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OR INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: EDITRA.034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 46866
LENGTH: 1148
TYPE: prt
ORGANISM: Bacillus anthracis
US-10-262-122A-46866

```

Query Match 50.9%; Score 174; DB 6; Length 1148;
 Best Local Similarity 53.0%; Pred. No. 2.2e+11;
 Matches 35; Conservative 13; Mismatches 18; Indels 0; Gaps 0;

RESULT 8


```
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 52088
LENGTH: 1144
TYPE: PRT
ORGANISM: Clostridium acetobutylicum
US-10-282-122A-52088

Query Match      48.8%; Score 167; DB 6; Length 1144;
Best Local Similarity 50.7%; Pred. No. 1.4e-10;
Matches 34; Conservative 11; Mismatches 22; Indels 0; Gaps 0;

QY 3 EIPAPLAGVSKILVKEGDTYKAGQTVLVEAMKMEFEINPTGQKVEKLVKERDAVQ 62
DB 1077 EIGASIGCNVYKVFYKQDKYKQDSLMYIEMKMEFNVSSEDTGVTGIVKEGDOYQ 1136
QY 63 GGGLIKI 69
DB 1137 GQLLVKL 1143

RESULT 12
US-10-282-122A-61290
Sequence 61290, Application US/10282122A
GENERAL INFORMATION:
APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
```

```
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 61290
LENGTH: 596
TYPE: PRT
ORGANISM: Legionella pneumophila
US-10-282-122A-61290

Query Match      48.5%; Score 166; DB 6; Length 596;
Best Local Similarity 51.5%; Pred. No. 7.8e-11;
Matches 35; Conservative 12; Mismatches 21; Indels 0; Gaps 0;

QY 2 GEIPAPLAGVSKILVKEGDTYKAGQTVLVEAMKMEFEINAPDQKVEKLVKERAVQ 61
DB 526 GDITVAIPGSIHAIHVSQDEVKAGQAVLVEAMKMEFEIKAPANGVAEILCKQKGYT 585
QY 62 GGGLIKI 69
DB 586 PGQVLIRV 593

RESULT 13
US-10-282-122A-60140
Sequence 60140, Application US/10282122A
GENERAL INFORMATION:
APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 60140
LENGTH: 596
TYPE: PRT
ORGANISM: Legionella pneumophila
US-10-282-122A-60140
```

```

; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 60140
; LENGTH: 588
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-10-282-122A-60140

Query Match
Best Local Similarity 48.2%; Score 165; DB 6; Length 588;
Matches 35; Conservative 9; Mismatches 22; Indels 0; Gaps 0;

QY 4 IPAPLAGTIVSKITVEGDPVKAGQTVLVLEAMKMETEINAPTDGKVEKVLKERDAVGG 63
DB 522 VTAPLAGTIVKVLASEGQVVAAGEVLLILEAMKMETEIRAAAGTVRGIAVAGDAVAVG 581
QY 64 GGLIKI 69
DB 582 DTLMTL 587

RESULT 14
US-10-282-122A-77079
; Sequence 77079, Application US/10282122A
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Tirawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 77079
; LENGTH: 597
; TYPE: PRT
```

```

; ORGANISM: Vibrio cholerae
US-10-282-122A-77079

Query Match
Best Local Similarity 48.0%; Score 164; DB 6; Length 597;
Matches 33; Conservative 14; Mismatches 19; Indels 0; Gaps 0;

QY 4 IPAPLAGTIVSKITVEGDPVKAGQTVLVLEAMKMETEINAPTDGKVEKVLKERDAVGG 63
DB 531 VAAPLAGTIVKIQVQGDVVAAGDVLLIVLEAMKMETEIRAAAGSVIOELHKGESVAVG 590
QY 64 GGLIKI 69
DB 591 ASLSTL 596

RESULT 15
US-10-282-122A-54468
; Sequence 54468, Application US/10282122A
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Tirawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 54468
; LENGTH: 599
; TYPE: PRT
; ORGANISM: Campylobacter jejuni
US-10-282-122A-54468

Query Match
Best Local Similarity 47.7%; Score 163; DB 6; Length 599;
Matches 33; Conservative 13; Mismatches 20; Indels 0; Gaps 0;

QY 1 BGEIPAPLAGTIVSKITVEGDPVKAGQTVLVLEAMKMETEINAPTDGKVEKVLKERDAV 60
DB 529 ENEVLAGTISGNFKITVNEGEVRSQAIWVLEAMKMETEIRANAPDGIITELCIKIGDTV 588
```


QY 61 QGCGCL 66
Db 589 NEGEVL 594

RESULT 16

US-10-224-539A-7
Sequence 7, Application US/10224539A
GENERAL INFORMATION:
APPLICANT: Tsai, Fong-Ying
TITLE OF INVENTION: 22325, A HUMAN BIOTIN-REQUIRING ENZYME
FILE REFERENCE: MP101-157P1M
CURRENT APPLICATION NUMBER: US/10/224,539A
PRIOR FILING DATE: 2002-08-20
PRIOR APPLICATION NUMBER: 60/313674
PRIOR FILING DATE: 2001-08-20
NUMBER OF SEQ ID NOS: 12
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 7
LENGTH: 75
TYPE: PRT
ORGANISM: Artificial Sequence
OTHER INFORMATION: consensus sequence
US-10-224-539A-7

Query Match

Best Local Similarity 46.9%; Score 160.5; DB 6; Length 75;
Matches 37; Conservative 11; Mismatches 19; Indels 7; Gaps 1;

QY 3 EIPAPLAG-----TVSKILVKGDTYKAGQTVLVEAMKETEINAPTDGKVEKVLVK 55
Db 2 EIKSPMIGSEYKEGTPVAEVLVKVDKRVKAGQVLCVEAMKEMEIPAPVAGVKEILVK 61

QY 56 ERDAVGGGGLIKI 69

Db 62 EGDIVEVGDPLAKI 75

RESULT 17

US-10-370-959-133
Sequence 133, Application US/10370959
GENERAL INFORMATION:
APPLICANT: Millennium Pharmaceuticals, Inc
APPLICANT: Meyers, Rachel E.
APPLICANT: Williamson, Mark W.
APPLICANT: Kapeller-Libermann, Rosana
APPLICANT: Macbeth, Kyle J.
APPLICANT: Hunter, John J.
APPLICANT: Rudolph-Owen, Laura A.
APPLICANT: Bandaru, Rajasekhar
TITLE OF INVENTION: NOVEL 13237, 18480, 2245, 16228, 7677,
TITLE OF INVENTION: 26320, 46619, 33166, 16836, 46867, 21617, 55562, 39228,
TITLE OF INVENTION: 62088, 46745, 23155, 21657, 42755, 32229, 22325, 46863 AND
FILE REFERENCE: MP103-0170NMIM
CURRENT APPLICATION NUMBER: US/10/370, 959
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: US 09/910,150
PRIOR FILING DATE: 2001-07-18
PRIOR APPLICATION NUMBER: US 60/219,028
PRIOR FILING DATE: 2000-07-18
PRIOR APPLICATION NUMBER: US 10/251,507
PRIOR FILING DATE: 2002-09-20
PRIOR APPLICATION NUMBER: US 09/715,479
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: US 60/218,053
PRIOR FILING DATE: 2000-07-13
PRIOR APPLICATION NUMBER: US 09/644,929
PRIOR FILING DATE: 2000-06-26
PRIOR APPLICATION NUMBER: US 09/775,117
PRIOR FILING DATE: 2001-02-01
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 136

PRIOR FILING DATE: 2000-06-16
PRIOR APPLICATION NUMBER: US 09/892,870
PRIOR FILING DATE: 2001-06-26
PRIOR APPLICATION NUMBER: US 60/214,174
PRIOR FILING DATE: 2000-06-26
PRIOR APPLICATION NUMBER: US 09/775,117
PRIOR FILING DATE: 2001-02-01
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 156
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 133
LENGTH: 75
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Consensus amino acid sequence
US-10-370-959-133

Query Match

Best Local Similarity 46.9%; Score 160.5; DB 6; Length 75;
Matches 37; Conservative 11; Mismatches 19; Indels 7; Gaps 1;

QY 3 EIPAPLAG-----TVSKILVKGDTYKAGQTVLVEAMKETEINAPTDGKVEKVLVK 55
Db 2 EIKSPMIGSEYKEGTPVAEVLVKVDKRVKAGQVLCVEAMKEMEIPAPVAGVKEILVK 61

QY 56 ERDAVGGGGLIKI 69

Db 62 EGDIVEVGDPLAKI 75

RESULT 18

US-10-377-097-133
Sequence 133, Application US/10377097
GENERAL INFORMATION:
APPLICANT: Millennium Pharmaceuticals, Inc
APPLICANT: Meyers, Rachel E.
APPLICANT: Williamson, Mark W.
APPLICANT: Kapeller-Libermann, Rosana
APPLICANT: Macbeth, Kyle J.
APPLICANT: Hunter, John J.
APPLICANT: Rudolph-Owen, Laura A.
APPLICANT: Bandaru, Rajasekhar
TITLE OF INVENTION: NOVEL 13237, 18480, 2245, 16228, 7677,
TITLE OF INVENTION: 26320, 46619, 33166, 16836, 46867, 21617, 55562, 39228,
TITLE OF INVENTION: 62088, 46745, 23155, 21657, 42755, 32229, 22325, 46863 AND
FILE REFERENCE: MP103-0350NMIM
CURRENT APPLICATION NUMBER: US/10/377,097
CURRENT FILING DATE: 2003-02-28
PRIOR APPLICATION NUMBER: US 09/910,150
PRIOR FILING DATE: 2001-07-18
PRIOR APPLICATION NUMBER: US 60/219,028
PRIOR FILING DATE: 2000-07-18
PRIOR APPLICATION NUMBER: US 10/251,507
PRIOR FILING DATE: 2002-09-20
PRIOR APPLICATION NUMBER: US 09/715,479
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: US 60/218,053
PRIOR FILING DATE: 2000-07-13
PRIOR APPLICATION NUMBER: US 09/644,929
PRIOR FILING DATE: 2000-08-23
PRIOR APPLICATION NUMBER: US 60/222,439
PRIOR FILING DATE: 2000-06-16
PRIOR APPLICATION NUMBER: US 09/892,870
PRIOR FILING DATE: 2001-06-26
PRIOR APPLICATION NUMBER: US 60/214,174
PRIOR FILING DATE: 2000-06-26
PRIOR APPLICATION NUMBER: US 09/775,117
PRIOR FILING DATE: 2001-02-01
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 136


```
Sequence 43666, Application US/10282122A
GENERAL INFORMATION:
APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remainder of prior application data removed - See file wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 43666
LENGTH: 607
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-10-282-122A-43666

Query Match
Best Local Similarity 46.2%; Score 158; DB 6; Length 607;
Matches 33; Conservative 12; Mismatches 23; Indels 0; Gaps 0;

QY 2 GEIPAPLAGTAKIIVKEDGTAKAGTAVLLEAMKMETEINAPTDGKVEKVLKRDVAQ 61
DB 538 GHVSTMPGNIVDLVYKEDSVKAGAVLITAMKMETEVQAGIAGTVKAIHVAGRDV 597
QY 62 GGQGLIKI 69
DB 598 PGEILIEI 605

RESULT 26
US-10-366-683-19134
GENERAL INFORMATION:
APPLICANT: Wang, Liangsu
APPLICANT: Rubenfield, Marc J.
APPLICANT: Nolling, Jork
APPLICANT: Delouhery, Craig
APPLICANT: Bush, David
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: PATH03-04
CURRENT APPLICATION NUMBER: US/10/366,683
CURRENT FILING DATE: 2003-02-13
```

```
PRIOR APPLICATION NUMBER: 09/252,991
PRIOR FILING DATE: 1999-02-18
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 19134
LENGTH: 612
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-10-366-683-19134

Query Match
Best Local Similarity 46.2%; Score 158; DB 6; Length 612;
Matches 33; Conservative 12; Mismatches 23; Indels 0; Gaps 0;

QY 2 GEIPAPLAGTAKIIVKEDGTAKAGTAVLLEAMKMETEINAPTDGKVEKVLKRDVAQ 61
DB 543 GHVSTMPGNIVDLVYKEDSVKAGAVLITAMKMETEVQAGIAGTVKAIHVAGRDV 602
QY 62 GGQGLIKI 69
DB 603 PGEILIEI 610

RESULT 27
US-10-419-128-19134
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/10/419,128
CURRENT FILING DATE: 2003-04-21
PRIOR APPLICATION NUMBER: US/09/252,991
PRIOR FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 19134
LENGTH: 612
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-10-419-128-19134

Query Match
Best Local Similarity 46.2%; Score 158; DB 6; Length 612;
Matches 33; Conservative 12; Mismatches 23; Indels 0; Gaps 0;

QY 2 GEIPAPLAGTAKIIVKEDGTAKAGTAVLLEAMKMETEINAPTDGKVEKVLKRDVAQ 61
DB 543 GHVSTMPGNIVDLVYKEDSVKAGAVLITAMKMETEVQAGIAGTVKAIHVAGRDV 602
QY 62 GGQGLIKI 69
DB 603 PGEILIEI 610

RESULT 28
US-10-282-122A-60488
GENERAL INFORMATION:
APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
```

APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining prior Application data removed - See file wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 60488
LENGTH: 1146
TYPE: PRT
ORGANISM: *Listeria monocytogenes*
US-10-282-122A-60488

Query Match
Best Local Similarity 43.9%; Score 157; DB 6; Length 1146;
Matches 29; Conservative 17; Mismatches 20; Indels 0; Gaps 0;

QY 4 IPAPLAGTVSKILVKEGDTVAKGQTVLVLEAMKMETEINAPTDGKVEKVLKERDAVGG 63
DB 1078 VGATMGVSQVGVVKKGDSYKKKDDPLITETAMKMETTIAQPFGEVSIVSDGTIESG 1137

QY 64 QGLIKI 69
DB 1138 DLIIEV 1143

RESULT 29
US-10-369-493-17338
Sequence 17338, Application US/10369493
GENERAL INFORMATION:
APPLICANT: Cao, Yongwei
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Goldman, Barry S.
APPLICANT: Chen, Xianteng
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
FILE REFERENCE: 38-10(52052)B
CURRENT APPLICATION NUMBER: US/10/369,493
CURRENT FILING DATE: 2003-02-28
PRIOR APPLICATION NUMBER: US 60/360,039
PRIOR FILING DATE: 2002-02-21
NUMBER OF SEQ ID NOS: 47374
SEQ ID NO 17338
LENGTH: 1150
TYPE: PRT
ORGANISM: *Bacillus halodurans*
US-10-369-493-17338

Query Match
Best Local Similarity 45.9%; Score 157; DB 6; Length 1150;
Matches 33; Conservative 13; Mismatches 20; Indels 0; Gaps 0;

QY 4 IPAPLAGTVSKILVKEGDTVAKGQTVLVLEAMKMETEINAPTDGKVEKVLKERDAVGG 63
DB 1082 IGASMGTVAKVLEKDGKQKGDHMITAMKMETTIAQPFGEVVALHVKGDAIQTG 1141
QY 64 QGLIKI 69
DB 1142 DLIIEV 1147

RESULT 30
US-10-282-122A-53885
Sequence 53885, Application US/10282122A
GENERAL INFORMATION:
APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining prior Application data removed - See file wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 53885
LENGTH: 1141
TYPE: PRT
ORGANISM: *Corynebacterium diptheriae*
US-10-282-122A-53885

Query Match
Best Local Similarity 45.5%; Score 155.5; DB 6; Length 1141;
Matches 32; Conservative 15; Mismatches 20; Indels 1; Gaps 1;

QY 2 GEIPPLAGTVSKILVKEGDTVAKGQTVLVLEAMKMETEINAPTDGKVEKVLKERDAVGG 61
DB 1074 GHVAAFPAGVYT-VTIEGATVAKGDPVAVIEMKMETTISATTDGTVNIVLQATKVE 1132
QY 62 GGOGLIKI 69
DB 1133 GGDLLIIV 1140

Search completed: May 1, 2003, 08:09:54

Thu May 1 08:55:12 2003

Job time : 235.156 secs

us-09-987-485-2.rapn

Page 12

GenCore version 5.1.4_p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 1, 2003, 07:49:49 ; Search time 99.5312 Seconds
(without alignments)
453.439 Million cell updates/sec

Title: US-09-987-485-2

Perfect score: 342
Sequence: 1 EGEIPAPLAGTVSKILVREG.....KVLVERDAVGGGGLIKIG 70

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 4569144 seqs, 644733110 residues

Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 100 summaries

Database :

Pending_Patents_AA_Main:*
1: /cgn2_6/ptodata/1/paa/PCRTUS_COMB.pep:*
2: /cgn2_6/ptodata/1/paa/US06_COMB.pep:*
3: /cgn2_6/ptodata/1/paa/US07_COMB.pep:*
4: /cgn2_6/ptodata/1/paa/US08_COMB.pep:*
5: /cgn2_6/ptodata/1/paa/US081_COMB.pep:*
6: /cgn2_6/ptodata/1/paa/US082_COMB.pep:*
7: /cgn2_6/ptodata/1/paa/US083_COMB.pep:*
8: /cgn2_6/ptodata/1/paa/US084_COMB.pep:*
9: /cgn2_6/ptodata/1/paa/US085_COMB.pep:*
10: /cgn2_6/ptodata/1/paa/US086_COMB.pep:*
11: /cgn2_6/ptodata/1/paa/US087_COMB.pep:*
12: /cgn2_6/ptodata/1/paa/US088_COMB.pep:*
13: /cgn2_6/ptodata/1/paa/US089_COMB.pep:*
14: /cgn2_6/ptodata/1/paa/US090_COMB.pep:*
15: /cgn2_6/ptodata/1/paa/US091_COMB.pep:*
16: /cgn2_6/ptodata/1/paa/US092_COMB.pep:*
17: /cgn2_6/ptodata/1/paa/US093_COMB.pep:*
18: /cgn2_6/ptodata/1/paa/US094_COMB.pep:*
19: /cgn2_6/ptodata/1/paa/US095_COMB.pep:*
20: /cgn2_6/ptodata/1/paa/US096_COMB.pep:*
21: /cgn2_6/ptodata/1/paa/US097_COMB.pep:*
22: /cgn2_6/ptodata/1/paa/US098_COMB.pep:*
23: /cgn2_6/ptodata/1/paa/US099_COMB.pep:*
24: /cgn2_6/ptodata/1/paa/US100_COMB.pep:*
25: /cgn2_6/ptodata/1/paa/US101_COMB.pep:*
26: /cgn2_6/ptodata/1/paa/US102_COMB.pep:*
27: /cgn2_6/ptodata/1/paa/US60_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	342	100.0	70	US-09-987-485-2	Sequence 2, Appl1
2	342	100.0	77	US-09-791-537-49119	Sequence 4, Appl1
3	342	100.0	122	US-09-987-485-1	Sequence 1, Appl1
4	342	100.0	123	US-09-791-537-132235	Sequence 132235
5	342	100.0	123	US-10-251-313-1	Sequence 1, Appl1
6	342	100.0	126	US-09-148-920-5	Sequence 5, Appl1

7	342	100.0	129	US-09-791-537-120787	Sequence 120787
8	342	100.0	133	US-10-251-313-7	Sequence 7, Appl1
9	342	100.0	256	US-09-148-920-3	Sequence 3, Appl1
10	342	100.0	266	US-09-148-920-7	Sequence 7, Appl1
11	185	57.0	145	US-09-791-537-111788	Sequence 111788
12	189	55.3	149	US-09-791-537-111800	Sequence 111800
13	180	52.6	571	US-09-791-537-31867	Sequence 31867, A
14	180	52.6	571	US-09-360-039-1263	Sequence 1263, Ap
15	180	52.6	571	US-09-360-039-20329	Sequence 20329, Ap
16	178	52.0	140	US-09-791-537-111780	Sequence 111780
17	174	50.9	655	US-09-791-537-31887	Sequence 31887, A
18	174	50.9	655	US-09-360-039-112	Sequence 112, App
19	173	50.6	984	US-09-791-537-24359	Sequence 24359, A
20	173	50.6	567	US-09-791-537-46258	Sequence 46258, A
21	173	50.6	567	US-09-360-039-1053	Sequence 1053, Ap
22	172	50.3	146	US-09-540-209B-9732	Sequence 9732, Ap
23	167	48.8	134	US-09-791-537-111798	Sequence 111798
24	167	48.8	596	US-09-791-537-107230	Sequence 107230
25	167	48.8	1076	US-09-360-039-13831	Sequence 13831, A
26	165	48.2	596	US-09-791-537-36890	Sequence 36890, A
27	163	47.7	142	US-09-791-537-111784	Sequence 111784
28	161	47.1	142	US-09-791-537-111784	Sequence 111784
29	159.5	46.6	574	US-09-360-039-14349	Sequence 14349, A
30	159.5	46.6	574	US-09-360-039-15017	Sequence 15017, A
31	159.5	46.6	575	US-09-360-039-11591	Sequence 11591, A
32	159.5	46.6	575	US-09-360-039-14608	Sequence 14608, A
33	159.5	46.6	576	US-09-739-449-11573	Sequence 11573, A
34	159.5	46.6	576	US-09-803-110-11573	Sequence 11573, A
35	158.5	46.3	1157	US-09-974-973-2	Sequence 974, A
36	158.5	46.3	1157	US-09-974-973-4	Sequence 974, A
37	158	46.2	612	US-09-252-991A-19134	Sequence 19134, A
38	157	45.9	1150	US-09-360-039-17338	Sequence 17338, A
39	156	45.6	169	US-09-902-540-16044	Sequence 16044, A
40	155.5	45.5	272	US-09-602-740-112	Sequence 112, App
41	155.5	45.5	272	US-09-602-740-114	Sequence 114, App
42	155.5	45.5	1140	US-09-529-043A-2	Sequence 2, Appl1
43	155.5	45.5	1140	US-09-529-043B-2	Sequence 2, Appl1
44	155.5	45.5	1140	US-09-738-626-4265	Sequence 4265, Ap
45	155.5	45.5	1140	US-09-791-537-45746	Sequence 45746, A
46	155.5	45.5	1140	US-09-974-973-19	Sequence 19, Appl
47	155.5	45.5	1140	US-10-045-072-2	Sequence 2, Appl1
48	153	44.7	620	US-09-791-537-31884	Sequence 31884, A
49	153	44.7	620	US-09-360-039-119	Sequence 119, App
50	152	44.4	167	US-09-791-537-63783	Sequence 63783, A
51	152	44.4	436	US-09-791-537-4143	Sequence 4143, Ap
52	152	44.4	591	US-09-791-537-121743	Sequence 121743
53	151.5	44.3	589	US-09-360-039-10079	Sequence 10079, A
54	150	43.9	180	US-09-107-532-5720	Sequence 5720, Ap
55	150	43.9	180	US-09-107-532A-5720	Sequence 5720, Ap
56	150	43.9	665	US-09-791-537-12352	Sequence 12352, Ap
57	148	43.3	1092	US-09-739-449-11312	Sequence 11312, A
58	148	43.3	1092	US-09-803-110-11312	Sequence 11312, A
59	148	43.3	1147	US-09-360-039-11450	Sequence 11450, A
60	148	43.3	1151	US-09-360-039-14817	Sequence 14817, A
61	148	43.3	1151	US-09-360-039-15006	Sequence 15006, A
62	146	42.7	69	US-09-897-516-8360	Sequence 8360, Ap
63	146	42.7	69	US-09-827-356-5175	Sequence 5175, Ap
64	146	42.7	989	US-08-827-356-5175	Sequence 8275, Ap
65	146	42.7	989	US-09-611-529-3973	Sequence 3973, Ap
66	146	42.7	1073	PCT-US02-03987-12361	Sequence 12361, A
67	146	42.7	1073	US-09-815-242-12361	Sequence 12361, A
68	146	42.7	1073	US-10-072-851-12361	Sequence 12361, A
69	146	42.7	1147	PCT-US02-03987-5468	Sequence 5468, Ap
70	146	42.7	1147	US-09-815-242-5468	Sequence 5468, Ap
71	146	42.7	1147	US-10-072-851-5468	Sequence 5468, Ap
72	146	42.7	1147	US-09-253-625-2442	Sequence 2442, Ap
73	146	42.7	1147	US-09-257-931-3012	Sequence 3012, Ap
74	146	42.7	1147	US-09-269-308-4003	Sequence 4003, Ap
75	146	42.7	1156	US-10-151-129-23	Sequence 23, Appl
76	146	42.7	1186	US-09-897-516-7481	Sequence 7481, Ap
77	146	42.7	1186	US-09-215-161-7481	Sequence 7481, Ap
78	145	42.4	130	US-09-791-537-132236	Sequence 132236
79	145	42.4	701	US-09-252-991A-27999	Sequence 27999, A

```
80 145 42.4 1147 21 US-09-791-537-57945 Sequence 57945, A
81 144 42.1 108 13 US-08-996-624-166 Sequence 166, App
82 144 42.1 108 15 US-09-156-181-166 Sequence 166, App
83 144 42.1 108 22 US-09-880-505-166 Sequence 166, App
84 144 42.1 108 24 US-10-051-643-166 Sequence 166, App
85 144 42.1 108 25 US-10-171-484-166 Sequence 166, App
86 144 42.1 147 15 US-09-134-000-6221 Sequence 6221, App
87 144 42.1 161 1 PCT-US02-03987-13363 Sequence 13363, A
88 144 42.1 161 1 PCT-US02-03987-13615 Sequence 13615, A
89 144 42.1 161 19 US-09-583-110-3907 Sequence 3907, App
90 144 42.1 161 21 US-09-752-069A-198 Sequence 198, App
91 144 42.1 161 22 US-09-815-242-13363 Sequence 13363, A
92 144 42.1 161 22 US-09-815-242-13363 Sequence 13363, A
93 144 42.1 161 24 US-10-072-851-13363 Sequence 13363, A
94 144 42.1 161 24 US-10-072-851-13615 Sequence 13615, A
95 144 42.1 161 27 PCT-US97-14436-402 Sequence 402, App
96 144 42.1 163 13 PCT-US97-14436-402 Sequence 402, App
97 144 42.1 163 13 US-08-911-503-402 Sequence 402, App
98 144 42.1 163 13 US-08-911-503A-402 Sequence 402, App
99 144 42.1 163 13 US-09-107-433-3683 Sequence 3683, App
100 144 42.1 243 15 US-09-156-181-199 Sequence 199, App
```

ALIGNMENTS

```
RESULT 1
; Sequence 2, Application US/09987485
; GENERAL INFORMATION:
; APPLICANT: Parrott, Michael
; TITLE OF INVENTION: Methods for the In Vivo Biotin Labeling of Polypeptides
; FILE REFERENCE: 15987/268653
; CURRENT APPLICATION NUMBER: US/09/987,485
; CURRENT FILING DATE: 2001-11-14
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 70
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-987-485-2
```

```
Query Match 100.0%; Score 342; DB 23; Length 70;
Best Local Similarity 100.0%; Pred. No. 2,7e-34;
Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 EGEIPAPLAGTIVSKILVEGDTVKAGQTVLLEAKMTEINAPTDGKVEKVLKERDAV 60
DB 1 EGEIPAPLAGTIVSKILVEGDTVKAGQTVLLEAKMTEINAPTDGKVEKVLKERDAV 60
QY 61 OGCGGLIKIG 70
DB 61 OGCGGLIKIG 70
```

```
RESULT 2
US-09-791-537-49119
; Sequence 49119, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Biomomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 49119
; LENGTH: 77
```

```
; TYPE: PRT
; ORGANISM: pbb 1DD2A
US-09-791-537-49119
Query Match 100.0%; Score 342; DB 21; Length 77;
Best Local Similarity 100.0%; Pred. No. 3.1e-34;
Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 EGEIPAPLAGTIVSKILVEGDTVKAGQTVLLEAKMTEINAPTDGKVEKVLKERDAV 60
DB 8 EGEIPAPLAGTIVSKILVEGDTVKAGQTVLLEAKMTEINAPTDGKVEKVLKERDAV 67
QY 61 OGCGGLIKIG 70
DB 68 OGCGGLIKIG 77
```

```
RESULT 3
US-09-987-485-1
; Sequence 1, Application US/09987485
; GENERAL INFORMATION:
; APPLICANT: Parrott, Michael
; TITLE OF INVENTION: Methods for the In Vivo Biotin Labeling of Polypeptides
; FILE REFERENCE: 15987/268653
; CURRENT APPLICATION NUMBER: US/09/987,485
; CURRENT FILING DATE: 2001-11-14
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 122
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-987-485-1
```

```
Query Match 100.0%; Score 342; DB 23; Length 122;
Best Local Similarity 100.0%; Pred. No. 5.9e-34;
Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 EGEIPAPLAGTIVSKILVEGDTVKAGQTVLLEAKMTEINAPTDGKVEKVLKERDAV 60
DB 53 EGEIPAPLAGTIVSKILVEGDTVKAGQTVLLEAKMTEINAPTDGKVEKVLKERDAV 112
QY 61 OGCGGLIKIG 70
DB 113 OGCGGLIKIG 122
```

```
RESULT 4
US-09-791-537-132235
; Sequence 132235, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Biomomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 132235
; LENGTH: 123
; TYPE: PRT
; ORGANISM: Propionibacterium freudenreichii subsp
US-09-791-537-132235
```

```
Query Match 100.0%; Score 342; DB 21; Length 123;
Best Local Similarity 100.0%; Pred. No. 5.9e-34;
Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 EGEIPAPLAGTIVSKILVEGDTVKAGQTVLLEAKMTEINAPTDGKVEKVLKERDAV 60
```


Db 54 EGEIPAPLAGTIVSKILVEGDTVAKAGOTVIVLEAMKMETEINAPTDGKVERVLYKERDAY 113
OY 61 QGGGGLIKIG 70
Db 114 QGGGGLIKIG 123

RESULT 5
US-10-251-313-1

; Sequence 1, Application US/10251313
; GENERAL INFORMATION:
; APPLICANT: AMBROSIOUS, DOROTHEE
; APPLICANT: LANZENDORFER, MARTIN
; APPLICANT: SCHRAEMT, MICHAEL
; APPLICANT: WATZELE, MANFRED
; TITLE OF INVENTION: IMPROVED METHOD FOR A SEQUENCE SPECIFIC BIOTINYLATION
; FILE REFERENCE: 506.1001
; CURRENT APPLICATION NUMBER: US/10/251.313
; CURRENT FILING DATE: 2002-09-20
; PRIOR APPLICATION NUMBER: EP 01129681.1
; PRIOR FILING DATE: 2001-12-13
; PRIOR APPLICATION NUMBER: EP 01122554.7
; PRIOR FILING DATE: 2001-09-25
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 123
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: 1.35
; OTHER INFORMATION: transcarboxylase subunit of Propionibacterium
; OTHER INFORMATION: sherman11
US-10-251-313-1

Query Match 100.0%; Score 342; DB 26; Length 123;
Best Local Similarity 100.0%; Pred. No. 5.9e-34;
Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EGEIPAPLAGTIVSKILVEGDTVAKAGOTVIVLEAMKMETEINAPTDGKVERVLYKERDAY 60
Db 54 EGEIPAPLAGTIVSKILVEGDTVAKAGOTVIVLEAMKMETEINAPTDGKVERVLYKERDAY 113
OY 61 QGGGGLIKIG 70
Db 114 QGGGGLIKIG 123

RESULT 6
US-09-148-920-5

; Sequence 5, Application US/09148920
; GENERAL INFORMATION:
; APPLICANT: Mullentix, Michael C.
; APPLICANT: Deutisch, John
; TITLE OF INVENTION: Recombinant Antigen Immunoassay for the Diagnosis of
; FILE REFERENCE: P-4131
; CURRENT APPLICATION NUMBER: US/09/148.920
; CURRENT FILING DATE: 1997-09-04
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 126
; TYPE: PRT
; ORGANISM: Treponema pallidum
; OTHER INFORMATION: Biotinylation peptide
US-09-148-920-5

Query Match 100.0%; Score 342; DB 15; Length 126;
Best Local Similarity 100.0%; Pred. No. 6.2e-34;
Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EGEIPAPLAGTIVSKILVEGDTVAKAGOTVIVLEAMKMETEINAPTDGKVERVLYKERDAY 60

Db 53 EGEIPAPLAGTIVSKILVEGDTVAKAGOTVIVLEAMKMETEINAPTDGKVERVLYKERDAY 112
OY 61 QGGGGLIKIG 70
Db 113 QGGGGLIKIG 122

RESULT 7
US-09-791-537-120787

; Sequence 120787, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791.537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 120787
; LENGTH: 129
; TYPE: PRT
; ORGANISM: Cloning vector pInPoint
US-09-791-537-120787

Query Match 100.0%; Score 342; DB 21; Length 129;
Best Local Similarity 100.0%; Pred. No. 6.4e-34;
Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EGEIPAPLAGTIVSKILVEGDTVAKAGOTVIVLEAMKMETEINAPTDGKVERVLYKERDAY 60
Db 53 EGEIPAPLAGTIVSKILVEGDTVAKAGOTVIVLEAMKMETEINAPTDGKVERVLYKERDAY 112
OY 61 QGGGGLIKIG 70
Db 113 QGGGGLIKIG 122

RESULT 8
US-10-251-313-7

; Sequence 7, Application US/10251313
; GENERAL INFORMATION:
; APPLICANT: AMBROSIOUS, DOROTHEE
; APPLICANT: LANZENDORFER, MARTIN
; APPLICANT: SCHRAEMT, MICHAEL
; APPLICANT: WATZELE, MANFRED
; TITLE OF INVENTION: IMPROVED METHOD FOR A SEQUENCE SPECIFIC BIOTINYLATION
; FILE REFERENCE: 506.1001
; CURRENT APPLICATION NUMBER: US/10/251.313
; CURRENT FILING DATE: 2002-09-20
; PRIOR APPLICATION NUMBER: EP 01129681.1
; PRIOR FILING DATE: 2001-12-13
; PRIOR APPLICATION NUMBER: EP 01122554.7
; PRIOR FILING DATE: 2001-09-25
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 133
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: pInPoint
; OTHER INFORMATION: Biotinylation peptide
US-10-251-313-7

Query Match 100.0%; Score 342; DB 26; Length 133;
Best Local Similarity 100.0%; Pred. No. 6.6e-34;
Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EGEIPAPLAGTIVSKILVEGDTVAKAGOTVIVLEAMKMETEINAPTDGKVERVLYKERDAY 60

```

Db 53 EGEIPAPLAGTYSKILVKEGDTVKAGQTVLVLEAMKMETEINAPTDGKVEKVLKERDAY 112
QY 61 QGGGGLIKIG 70
Db 113 QGGGGLIKIG 122

RESULT 9
US-09-148-920-3
; Sequence 3, Application US/09148920
; GENERAL INFORMATION:
; APPLICANT: Mullenix, Michael C.
; APPLICANT: Deutsch, John
; TITLE OF INVENTION: Recombinant Antigen Immunoassay for the Diagnosis of
; FILE REFERENCE: P-4131
; CURRENT APPLICATION NUMBER: US/09/148,920
; CURRENT FILING DATE: 1997-09-04
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 256
; TYPE: PRT
; ORGANISM: Treponema pallidum
US-09-148-920-3

Query Match
Best Local Similarity 100.0%; Score 342; DB 15; Length 256;
Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EGEIPAPLAGTYSKILVKEGDTVKAGQTVLVLEAMKMETEINAPTDGKVEKVLKERDAY 60
Db 53 EGEIPAPLAGTYSKILVKEGDTVKAGQTVLVLEAMKMETEINAPTDGKVEKVLKERDAY 112
QY 61 QGGGGLIKIG 70
Db 113 QGGGGLIKIG 122

RESULT 10
US-09-148-920-7
; Sequence 7, Application US/09148920
; GENERAL INFORMATION:
; APPLICANT: Mullenix, Michael C.
; APPLICANT: Deutsch, John
; TITLE OF INVENTION: Recombinant Antigen Immunoassay for the Diagnosis of
; FILE REFERENCE: P-4131
; CURRENT APPLICATION NUMBER: US/09/148,920
; CURRENT FILING DATE: 1997-09-04
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 266
; TYPE: PRT
; ORGANISM: Treponema pallidum
US-09-148-920-7

Query Match
Best Local Similarity 100.0%; Score 342; DB 15; Length 266;
Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EGEIPAPLAGTYSKILVKEGDTVKAGQTVLVLEAMKMETEINAPTDGKVEKVLKERDAY 60
Db 53 EGEIPAPLAGTYSKILVKEGDTVKAGQTVLVLEAMKMETEINAPTDGKVEKVLKERDAY 112
QY 61 QGGGGLIKIG 70
Db 113 QGGGGLIKIG 122

RESULT 11
US-09-791-537-111788
; Sequence 111788, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Blomomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 111788
; LENGTH: 145
; TYPE: PRT
; ORGANISM: Pyrococcus abyssi
US-09-791-537-111788

Query Match
Best Local Similarity 57.0%; Score 195; DB 21; Length 145;
Matches 41; Conservative 10; Mismatches 19; Indels 0; Gaps 0;

QY 1 EGEIPAPLAGTYSKILVKEGDTVKAGQTVLVLEAMKMETEINAPTDGKVEKVLKERDAY 60
Db 76 ENVVAPMPKVKILVQEGQGVKLGGLLLEAMKMETEINAPTDGKVEKVLKERDAY 135
QY 61 QGGGGLIKIG 70
Db 136 DTGQPLIEIG 145

RESULT 12
US-09-791-537-111800
; Sequence 111800, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Blomomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 111800
; LENGTH: 149
; TYPE: PRT
; ORGANISM: Pyrococcus horikoshii
US-09-791-537-111800

Query Match
Best Local Similarity 55.3%; Score 189; DB 21; Length 149;
Matches 38; Conservative 14; Mismatches 18; Indels 0; Gaps 0;

QY 1 EGEIPAPLAGTYSKILVKEGDTVKAGQTVLVLEAMKMETEINAPTDGKVEKVLKERDAY 60
Db 80 ENVVAPMPKVKILVQEGQGVKLGGLLLEAMKMETEINAPTDGKVEKVLKERDAY 139
QY 61 QGGGGLIKIG 70
Db 140 DTGQPLIEIG 149

RESULT 13
US-09-791-537-31867
; Sequence 31867, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Blomomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
```

;; TITLE OF INVENTION: METHODS OF USE THEREOF
;; FILE REFERENCE: 261/210
;; CURRENT APPLICATION NUMBER: US/09/791,537
;; CURRENT FILING DATE: 2001-02-22
;; NUMBER OF SEQ ID NOS: 153055
;; SOFTWARE: PatentIn version 3.0
;; SEQ ID NO 31867
;; LENGTH: 571
;; TYPE: PRF
;; ORGANISM: Pyrococcus horikoshii
US-09-791-537-31867

Query Match
Best Local Similarity 54.5%; Score 180; DB 21; Length 571;
Pred. No. 6.8e-13;
Matches 36; Conservative 14; Mismatches 16; Indels 0; Gaps 0;

QY 4 IPAPLAGTAVSKILVKEGDTYKAGTAVLVLEAMKMETEINAPTDGKVEKVLKEDAVOG 63
DB 505 VSAPMPGKVLRLVGVDRVVGGLLVLEAMKMETEINPSRDSGVVRLVKEGEAVDTG 564
QY 64 QGLIKI 69
DB 565 QPLIEL 570

RESULT 14
US-60-360-039-1263
;; Sequence 1263, Application US/60360039
;; GENERAL INFORMATION:
;; APPLICANT: Cao, Yongwei
;; APPLICANT: Chen, Xianfeng
;; APPLICANT: Goldman, Barry S.
;; APPLICANT: Hinkle, Gregory J.
;; APPLICANT: Slater, Steven C.
;; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
;; FILE REFERENCE: 38-10(52052)A
;; CURRENT APPLICATION NUMBER: US/60/360,039
;; CURRENT FILING DATE: 2002-02-21
;; NUMBER OF SEQ ID NOS: 47374
;; SEQ ID NO 1263
;; LENGTH: 571
;; TYPE: PRF
;; ORGANISM: Pyrococcus horikoshii
US-60-360-039-1263

Query Match
Best Local Similarity 52.6%; Score 180; DB 27; Length 571;
Pred. No. 6.8e-13;
Matches 36; Conservative 14; Mismatches 16; Indels 0; Gaps 0;

QY 4 IPAPLAGTAVSKILVKEGDTYKAGTAVLVLEAMKMETEINAPTDGKVEKVLKEDAVOG 63
DB 505 VSAPMPGKVLRLVGVDRVVGGLLVLEAMKMETEINPSRDSGVVRLVKEGEAVDTG 564
QY 64 QGLIKI 69
DB 565 QPLIEL 570

RESULT 15
US-60-360-039-20329
;; Sequence 20329, Application US/60360039
;; GENERAL INFORMATION:
;; APPLICANT: Cao, Yongwei
;; APPLICANT: Chen, Xianfeng
;; APPLICANT: Goldman, Barry S.
;; APPLICANT: Hinkle, Gregory J.
;; APPLICANT: Slater, Steven C.
;; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
;; FILE REFERENCE: 38-10(52052)A
;; CURRENT APPLICATION NUMBER: US/60/360,039
;; CURRENT FILING DATE: 2002-02-21

;; NUMBER OF SEQ ID NOS: 47374
;; SEQ ID NO 20329
;; LENGTH: 571
;; TYPE: PRF
;; ORGANISM: Pyrococcus horikoshii
US-60-360-039-20329

Query Match
Best Local Similarity 52.6%; Score 180; DB 27; Length 571;
Pred. No. 6.8e-13;
Matches 36; Conservative 14; Mismatches 16; Indels 0; Gaps 0;

QY 4 IPAPLAGTAVSKILVKEGDTYKAGTAVLVLEAMKMETEINAPTDGKVEKVLKEDAVOG 63
DB 505 VSAPMPGKVLRLVGVDRVVGGLLVLEAMKMETEINPSRDSGVVRLVKEGEAVDTG 564
QY 64 QGLIKI 69
DB 565 QPLIEL 570

RESULT 16
US-09-791-537-111780
;; Sequence 111780, Application US/09791537
;; GENERAL INFORMATION:
;; APPLICANT: Biomolix, Inc.
;; APPLICANT: Danzer, Derek
;; APPLICANT: Danzer, Joseph
;; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
;; FILE REFERENCE: 261/210
;; CURRENT APPLICATION NUMBER: US/09/791,537
;; CURRENT FILING DATE: 2001-02-22
;; NUMBER OF SEQ ID NOS: 153055
;; SOFTWARE: PatentIn version 3.0
;; SEQ ID NO 111780
;; LENGTH: 140
;; TYPE: PRF
;; ORGANISM: Archaeoglobus fulgidus
US-09-791-537-111780

Query Match
Best Local Similarity 52.0%; Score 178; DB 21; Length 140;
Pred. No. 1.7e-13;
Matches 35; Conservative 13; Mismatches 16; Indels 0; Gaps 0;

QY 4 IPAPLAGTAVSKILVKEGDTYKAGTAVLVLEAMKMETEINAPTDGKVEKVLKEDAVOG 63
DB 74 ITAPMACVTKILKKVGEKAKAGETVLIIEMKMETEINPSRDSGVVRLVKEGEAVDTG 133
QY 64 QGLI 67
DB 134 DVLV 137

RESULT 17
US-09-791-537-31887
;; Sequence 31887, Application US/09791537
;; GENERAL INFORMATION:
;; APPLICANT: Biomolix, Inc.
;; APPLICANT: Debe, Derek
;; APPLICANT: Danzer, Joseph
;; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
;; FILE REFERENCE: 261/210
;; CURRENT APPLICATION NUMBER: US/09/791,537
;; CURRENT FILING DATE: 2001-02-22
;; NUMBER OF SEQ ID NOS: 153055
;; SOFTWARE: PatentIn version 3.0
;; SEQ ID NO 31887
;; LENGTH: 655
;; TYPE: PRF
;; ORGANISM: Aquifex aeolicus
US-09-791-537-31887

Query Match 50.9%; Score 174; DB 21; Length 655;
Best Local Similarity 50.0%; Pred. No. 4, 6e-12;
Matches 34; Conservative 13; Mismatches 21; Indels 0; Gaps 0;

QY 2 GEIPAPLAGTYSKILVKEGDTYKAGOTVLEAMKMETEINAPTDGKVEKVLKERDVAQ 61
DB 578 GDTTPMPKGVKILVKEGEPVQOGOTVATVEAMKMEINAHIDGIIVKIFAKPDQVN 637
QY 62 GGQGLIKI 69
DB 638 PDQAIMRI 645

RESULT 18
US-60-360-039-112
; Sequence 112, Application US/60360039
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Chen, Xianfeng
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)A
; CURRENT APPLICATION NUMBER: US/60/360, 039
; CURRENT FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 112
; LENGTH: 655
; TYPE: PRT
; ORGANISM: Aquifex aeolicus
US-60-360-039-112

Query Match 50.9%; Score 174; DB 27; Length 655;
Best Local Similarity 50.0%; Pred. No. 4, 6e-12;
Matches 34; Conservative 13; Mismatches 21; Indels 0; Gaps 0;

QY 2 GEIPAPLAGTYSKILVKEGDTYKAGOTVLEAMKMETEINAPTDGKVEKVLKERDVAQ 61
DB 578 GDTTPMPKGVKILVKEGEPVQOGOTVATVEAMKMEINAHIDGIIVKIFAKPDQVN 637
QY 62 GGQGLIKI 69
DB 638 PDQAIMRI 645

RESULT 19
US-09-791-537-24359
; Sequence 24359, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Biomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 24359
; LENGTH: 984
; TYPE: PRT
; ORGANISM: Bacillus cereus
US-09-791-537-24359

Query Match 50.9%; Score 174; DB 21; Length 984;
Best Local Similarity 53.0%; Pred. No. 8, 1e-12;
Matches 35; Conservative 13; Mismatches 18; Indels 0; Gaps 0;

QY 4 IPAPLAGTYSKILVKEGDTYKAGOTVLEAMKMETEINAPTDGKVEKVLKERDVAQ 63
DB 578 GDTTPMPKGVKILVKEGEPVQOGOTVATVEAMKMEINAHIDGIIVKIFAKPDQVN 637

DB 917 ISATMPGVYIKVYKGEDEKKGDSNAITEAMKMETTYQAPENGAKKYVNDGAIQTG 976
QY 64 OGGLIKI 69
DB 977 DLLIEL 982

RESULT 20
US-09-791-537-46258
; Sequence 46258, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Biomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 46258
; LENGTH: 567
; TYPE: PRT
; ORGANISM: Methanococcus jannaschii
US-09-791-537-46258

Query Match 50.6%; Score 173; DB 21; Length 567;
Best Local Similarity 50.7%; Pred. No. 5e-12;
Matches 35; Conservative 12; Mismatches 22; Indels 0; Gaps 0;

QY 1 EGEIPAPLAGTYSKILVKEGDTYKAGOTVLEAMKMETEINAPTDGKVEKVLKERDVAQ 60
DB 498 EGAVTSPFRGAVTKIKVKEGDKVKKGDVIVLEAMKMEHPLESPEGVETRIIDEGAV 557
QY 61 OGQGLIKI 69
DB 558 NVGDVIMIT 566

RESULT 21
US-60-360-039-1053
; Sequence 1053, Application US/60360039
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Chen, Xianfeng
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)A
; CURRENT APPLICATION NUMBER: US/60/360, 039
; CURRENT FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 1053
; LENGTH: 567
; TYPE: PRT
; ORGANISM: Methanococcus jannaschii
US-60-360-039-1053

Query Match 50.6%; Score 173; DB 27; Length 567;
Best Local Similarity 50.7%; Pred. No. 5e-12;
Matches 35; Conservative 12; Mismatches 22; Indels 0; Gaps 0;

QY 1 EGEIPAPLAGTYSKILVKEGDTYKAGOTVLEAMKMETEINAPTDGKVEKVLKERDVAQ 60
DB 498 EGAVTSPFRGAVTKIKVKEGDKVKKGDVIVLEAMKMEHPLESPEGVETRIIDEGAV 557
QY 61 OGQGLIKI 69
DB 558 NVGDVIMIT 566

```
RESULT 22
US-09-540-209B-9732
; Sequence 9732, Application US/09540209B
; GENERAL INFORMATION:
; APPLICANT: Gary L. Bretton
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO BACTERIOIDES FR
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.1001-001
; CURRENT APPLICATION NUMBER: US/09/540.209B
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 10444
; SEQ ID NO 9732
; LENGTH: 146
; TYPE: PR
; ORGANISM: B.fragilis
US-09-540-209B-9732

Query Match
Best Local Similarity 50.3%; Score 172; DB 19; Length 146;
Best Local Similarity 53.0%; Pred. No. 9.8e-13;
Matches 35; Conservative 11; Mismatches 20; Indels 0; Gaps 0;

QY 4 IPAPLAGTIVSKILVKGDTVKAQGVTLVLEAMKMETEINAPTDGKVEKYLKERDAVQGG 63
Db 80 VKSPLPVYIILDKYKESGDTVRGQITITLLEAMKKNINAKDKVAETIKYKGDVLEGG 139
QY 64 QGLIKI 69
Db 140 TDLVIT 145

RESULT 23
US-09-791-537-111798
; Sequence 111798, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Biomix, Inc.
; APPLICANT: Debe, Derek
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791.537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 111798
; LENGTH: 134
; TYPE: PR
; ORGANISM: Thermotoga maritima
US-09-791-537-111798

Query Match
Best Local Similarity 48.8%; Score 167; DB 21; Length 134;
Best Local Similarity 51.5%; Pred. No. 3.6e-12;
Matches 34; Conservative 12; Mismatches 20; Indels 0; Gaps 0;

QY 4 IPAPLAGTIVSKILVKGDTVKAQGVTLVLEAMKMETEINAPTDGKVEKYLKERDAVQGG 63
Db 68 VKAPAGIVLKVLEKGVKVVNDKLIVLEAMKMETEINAPTDGKVEKYLKERDAVQGG 127
QY 64 QGLIKI 69
Db 128 QILMKI 133

RESULT 24
US-09-791-537-107230
; Sequence 107230, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Biomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
; TITLE OF INVENTION: METHODS OF USE THEREOF
```

```
FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791.537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 107230
; LENGTH: 596
; TYPE: PR
; ORGANISM: Legionella pneumophila
US-09-791-537-107230

Query Match
Best Local Similarity 48.8%; Score 167; DB 21; Length 596;
Best Local Similarity 51.5%; Pred. No. 3e-11;
Matches 35; Conservative 12; Mismatches 21; Indels 0; Gaps 0;

QY 2 GEIPAPLAGTIVSKILVKGDTVKAQGVTLVLEAMKMETEINAPTDGKVEKYLKERDAVQ 61
Db 526 GDITVALPGSLIAIHVSAGDEVKAGQAVLVLEAMKMETEINAPANGVAETILCKGDKVT 585
QY 62 GGQGLIKI 69
Db 586 PGVLIIV 593

RESULT 25
US-60-360-039-13831
; Sequence 13831, Application US/60360039
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Chen, Xianfeng
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-1052052A
; CURRENT APPLICATION NUMBER: US/60/360.039
; CURRENT FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 13831
; LENGTH: 1076
; TYPE: PR
; ORGANISM: Pseudomonas fluorescens
US-60-360-039-13831

Query Match
Best Local Similarity 48.8%; Score 167; DB 27; Length 1076;
Best Local Similarity 51.5%; Pred. No. 6.8e-11;
Matches 35; Conservative 10; Mismatches 23; Indels 0; Gaps 0;

QY 2 GEIPAPLAGTIVSKILVKGDTVKAQGVTLVLEAMKMETEINAPTDGKVEKYLKERDAVQ 61
Db 1008 GHVSTPMGNIVDVLYKESGDTVRKAGQAVLVLEAMKMETEINAPANGVAETILCKGDKVT 1067
QY 62 GGQGLIKI 69
Db 1068 PGELIIEI 1075

RESULT 26
US-09-791-537-72996
; Sequence 72996, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Biomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791.537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 72996
```

LENGTH: 596
TYPE: PRT
ORGANISM: Klebsiella pneumoniae
US-09-791-537-72996

Query Match 48.2%; Score 165; DB 21; Length 596;
Best Local Similarity 53.0%; Pred. No. 5.2e-11;
Matches 35; Conservative 9; Mismatches 22; Indels 0; Gaps 0;

Db 530 VMAPLACTIKVLAISGQVYAAEVLIIIEAMKMEIETIRAAAGYVAGIYKAGDAVAVG 589

QY 64 QGIRIKI 69
Db 590 DFLMTL 595

RESULT 27
US-09-791-537-36890
Sequence 36890, Application US/09791537
GENERAL INFORMATION:
APPLICANT: Bionomix, Inc.
APPLICANT: Danzer, Derek
TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBER
FILE REFERENCE: 261/210
CURRENT APPLICATION NUMBER: US/09/791,537
CURRENT FILING DATE: 2001-02-22
NUMBER OF SEQ ID NOS: 153055
SOFTWARE: PatentIn version 3.0
SEQ ID NO 36890
LENGTH: 599
TYPE: PRT
ORGANISM: Campylobacter jejuni
US-09-791-537-36890

Query Match 47.7%; Score 163; DB 21; Length 599;
Best Local Similarity 50.0%; Pred. No. 9.4e-11;
Matches 33; Conservative 13; Mismatches 20; Indels 0; Gaps 0;

QY 1 EGEIPAPLAGVSKILVKGEDTVKAGQTVLVLEAMKMEIETINAPTDGKVEKVLKERVAV 60
Db 529 ENEVLAGISGNVFKIYNEGEVKSQDAIVLEAMKMEIETINAPTDGKVEKVLKERVAV 588

QY 61 QCGGGL 66
Db 589 NEGEVL 594

RESULT 28
US-09-791-537-111784
Sequence 111784, Application US/09791537
GENERAL INFORMATION:
APPLICANT: Bionomix, Inc.
APPLICANT: Debe, Derek
TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBER
FILE REFERENCE: 261/210
CURRENT APPLICATION NUMBER: US/09/791,537
CURRENT FILING DATE: 2001-02-22
NUMBER OF SEQ ID NOS: 153055
SOFTWARE: PatentIn version 3.0
SEQ ID NO 111784
LENGTH: 142
TYPE: PRT
ORGANISM: Archaeoglobus fulgidus
US-09-791-537-111784

Query Match 47.1%; Score 161; DB 21; Length 142;
Best Local Similarity 50.8%; Pred. No. 2.2e-11;

Matches 32; Conservative 13; Mismatches 18; Indels 0; Gaps 0;

QY 1 EGEIPAPLAGVSKILVKGEDTVKAGQTVLVLEAMKMEIETINAPTDGKVEKVLKERVAV 60
Db 70 ENAVTSMIPGVKILVKGDKVAGEPVIVESKMEIETISPEGVAVETLVKRGORI 129

QY 61 QCG 63
Db 130 EAG 132

RESULT 29
US-60-360-039-14349
Sequence 14349, Application US/60360039
GENERAL INFORMATION:
APPLICANT: Cao, Yongwei
APPLICANT: Chen, Xianfeng
APPLICANT: Goldman, Barry S.
APPLICANT: Hinkle, Gregory J.
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
FILE REFERENCE: 38-10(52052)A
CURRENT APPLICATION NUMBER: US/60/360,039
CURRENT FILING DATE: 2002-02-21
NUMBER OF SEQ ID NOS: 47374
SEQ ID NO 14349
LENGTH: 574
TYPE: PRT
ORGANISM: Agrobacterium tumefaciens
US-60-360-039-14349

Query Match 46.6%; Score 159.5; DB 27; Length 574;
Best Local Similarity 50.7%; Pred. No. 2.4e-10;
Matches 35; Conservative 13; Mismatches 20; Indels 1; Gaps 1;

QY 1 EGEIPAPLAGVSKILVKGEDTVKAGQTVLVLEAMKMEIETINAPTDGKVEKVLKERVAV 60
Db 507 EGEIPAPVSGTLOSFVKDGETVSEGDLLAVEMAKMETQIVATRAKV-RLIVREGDYL 565

QY 61 QCGGGLIKI 69
Db 566 QAGATLIDI 574

RESULT 30
US-60-360-039-15017
Sequence 15017, Application US/60360039
GENERAL INFORMATION:
APPLICANT: Cao, Yongwei
APPLICANT: Chen, Xianfeng
APPLICANT: Goldman, Barry S.
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
FILE REFERENCE: 38-10(52052)A
CURRENT APPLICATION NUMBER: US/60/360,039
CURRENT FILING DATE: 2002-02-21
NUMBER OF SEQ ID NOS: 47374
SEQ ID NO 15017
LENGTH: 574
TYPE: PRT
ORGANISM: Agrobacterium tumefaciens
US-60-360-039-15017

Query Match 46.6%; Score 159.5; DB 27; Length 574;
Best Local Similarity 50.7%; Pred. No. 2.4e-10;
Matches 35; Conservative 13; Mismatches 20; Indels 1; Gaps 1;

QY 1 EGEIPAPLAGVSKILVKGEDTVKAGQTVLVLEAMKMEIETINAPTDGKVEKVLKERVAV 60
Db 507 EGEIPAPVSGTLOSFVKDGETVSEGDLLAVEMAKMETQIVATRAKV-RLIVREGDYL 565

QY 61 QGGGULIKI 69
| | | | |
Db 566 QAGATLIDI 574

Search completed: May 1, 2003, 07:58:37
Job time : 100.531 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 1, 2003, 07:52:52 ; Search time 134.531 Seconds

(without alignments)
44.897 Million cell updates/sec

Title: US-09-987-485-2

Perfect score: 342
Sequence: 1 EGRIPLAGTYSKIIVKES.....KVLKERDAVGGGLIKIG 70Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328255 seqs, 86286685 residues

Total number of hits satisfying chosen parameters: 328255

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database :

Published Applications AA:*

1: /cgn2_6/ptodata/2/pubppaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/2/pubppaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubppaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubppaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubppaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubppaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubppaa/PCTUS_PUBCOMB.pep:*
8: /cgn2_6/ptodata/2/pubppaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubppaa/US09_NEW_PUB.pep:*
10: /cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubppaa/US10_NEW_PUB.pep:*
12: /cgn2_6/ptodata/2/pubppaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/2/pubppaa/US60_NEW_PUB.pep:*
14: /cgn2_6/ptodata/2/pubppaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Length	DB ID	Description
1	342	100.0	70	US-09-987-485-2	Sequence 2, Appl1
2	342	100.0	122	US-09-987-485-1	Sequence 1, Appl1
3	158.5	46.3	1157	US-09-974-973-2	Sequence 4, Appl1
4	158.5	46.3	1157	US-09-974-973-4	Sequence 2, Appl1
5	155.5	45.5	1140	US-09-974-973-19	Sequence 19, Appl1
6	155.5	45.5	1140	US-09-738-626-4265	Sequence 4265, Ap
7	155.5	45.5	1140	US-10-045-072-2	Sequence 2, Appl1
8	147	43.0	134	US-10-169-048-16	Sequence 16, Appl1
9	146	42.7	1073	US-09-815-242-12361	Sequence 12361, A
10	146	42.7	1147	US-09-815-242-5468	Sequence 5468, Ap
11	144	42.1	108	US-10-051-643-166	Sequence 166, App
12	144	42.1	108	US-09-880-505-166	Sequence 166, App
13	144	42.1	161	US-09-815-242-13363	Sequence 13363, A
14	144	42.1	161	US-09-815-242-13615	Sequence 13615, A
15	144	42.1	243	US-10-051-643-199	Sequence 199, App
16	144	42.1	1142	US-09-815-242-10806	Sequence 10806, A
17	143	40.8	725	US-10-160-501-17	Sequence 17, Appl
18	139.5	40.8	160	US-09-815-242-4962	Sequence 4962, Ap
19	139.5	40.8	162	US-09-815-242-10922	Sequence 10922, A

20	126.5	37.0	182	US-09-767-479-2	Sequence 2, Appl1
21	126	36.9	591	US-09-738-626-6940	Sequence 6940, Ap
22	119.5	34.9	154	US-09-815-242-12581	Sequence 12581, A
23	119.5	34.9	158	US-09-767-479-4	Sequence 4, Appl1
24	115	33.6	71	US-09-767-479-1	Sequence 1, Appl1
25	113.5	33.2	630	US-09-815-242-10036	Sequence 10036, A
26	112.5	30.0	155	US-09-815-242-11159	Sequence 11159, A
27	101.5	29.7	156	US-09-815-242-11322	Sequence 11322, A
28	101.5	29.7	156	US-10-081-051-16	Sequence 16, Appl
29	99	28.9	675	US-09-738-626-5917	Sequence 5917, Ap
30	98.5	28.8	156	US-09-815-242-12062	Sequence 12062, A
31	97.5	28.5	71	US-09-738-626-4965	Sequence 4965, Ap
32	97.5	28.5	156	US-09-815-242-10329	Sequence 10329, A
33	95.5	27.9	156	US-09-815-242-13886	Sequence 13886, A
34	88	25.7	430	US-09-815-242-5802	Sequence 5802, Ap
35	88	25.7	430	US-09-815-242-12935	Sequence 12935, A
36	87	25.4	50	US-09-791-171-129	Sequence 129, App
37	87	25.4	539	US-09-815-242-10497	Sequence 10497, A
38	80	23.4	50	US-09-791-171-81	Sequence 81, Appl
39	80	23.4	50	US-09-791-171-127	Sequence 127, App
40	77.5	22.7	409	US-09-815-242-11270	Sequence 11270, A
41	76.5	22.4	424	US-09-815-242-12587	Sequence 12587, A
42	76.5	22.4	425	US-09-815-242-5485	Sequence 5485, Ap
43	76.5	22.4	819	US-09-820-863A-62	Sequence 62, Appl
44	72.5	21.2	407	US-10-078-107-2	Sequence 2, Appl1
45	72.5	21.2	407	US-10-077-751-2	Sequence 4, Appl1
46	72.5	21.2	407	US-09-784-208-4	Sequence 4, Appl1
47	71.5	20.9	49	US-09-791-171-128	Sequence 128, App
48	71.5	20.9	329	US-08-945-038-4	Sequence 4, Appl1
49	71.5	20.9	355	US-09-741-669-435	Sequence 435, App
50	70	20.5	475	US-09-884-666-37	Sequence 37, Appl
51	69.5	20.3	215	US-09-815-242-12956	Sequence 12956, A
52	69.5	20.3	215	US-09-815-242-13084	Sequence 13084, A
53	66.5	19.4	1525	US-09-782-114-1	Sequence 1, Appl1
54	65	18.9	432	US-09-815-242-10693	Sequence 10693, A
55	64.5	18.9	107	US-09-815-242-14054	Sequence 14054, A
56	63	18.4	107	US-09-815-242-10831	Sequence 10831, A
57	63	18.4	129	US-09-815-242-10289	Sequence 10289, A
58	63	18.4	129	US-09-815-242-13879	Sequence 13879, A
59	62.5	18.3	405	US-10-077-751-6	Sequence 6, Appl1
60	62.5	18.3	405	US-09-815-242-10096	Sequence 10096, A
61	62.5	18.3	142	US-09-815-242-5718	Sequence 5718, Ap
62	62	18.1	280	US-09-815-242-12479	Sequence 12479, A
63	62	18.1	78	US-09-841-132-435	Sequence 435, App
64	61	17.8	474	US-09-738-626-8856	Sequence 656, App
65	61	17.8	478	US-09-884-666-40	Sequence 40, Appl
66	61	17.8	135	US-09-815-242-5842	Sequence 5842, Ap
67	60	17.5	263	US-09-815-242-13018	Sequence 13018, A
68	60	17.5	263	US-09-815-242-13152	Sequence 13152, A
69	60	17.5	285	US-09-815-242-11913	Sequence 11913, A
70	60	17.5	951	US-09-815-242-13853	Sequence 13853, A
71	59	17.3	561	US-10-078-770-154	Sequence 154, App
72	58.5	17.1	1702	US-09-839-996-5	Sequence 5, Appl1
73	58.5	17.1	1702	US-10-080-305-5	Sequence 26, Appl1
74	58.5	17.1	98	US-10-078-929-26	Sequence 13614, A
75	58	17.0	455	US-09-815-242-13364	Sequence 13364, A
76	58	17.0	455	US-09-815-242-13617	Sequence 13617, A
77	58	17.0	455	US-09-815-242-13617	Sequence 39, Appl
78	57.5	16.8	478	US-09-884-666-39	Sequence 721, App
79	57	16.7	523	US-10-102-806-721	Sequence 6486, App
80	57	16.7	548	US-09-738-626-6486	Sequence 38018, A
81	57	16.7	567	US-09-815-242-13563	Sequence 13563, A
82	57	16.7	597	US-09-864-761-38018	Sequence 13618, A
83	57	16.7	687	US-09-815-242-5822	Sequence 5822, Ap
84	57	16.7	719	US-09-815-242-12942	Sequence 12942, A
85	57	16.7	951	US-09-815-242-10465	Sequence 10465, A
86	57	16.7	2910	US-10-124-800-2	Sequence 2, Appl1
87	57	16.7	26926	US-09-759-5088-2	Sequence 2, Appl1
88	56.5	16.5	451	US-09-815-242-5633	Sequence 5633, Ap
89	56.5	16.5	451	US-09-815-242-12384	Sequence 12384, A
90	56.5	16.5	452	US-10-084-205-15	Sequence 15, Appl
91	56.5	16.5	452	US-09-925-637-16	Sequence 16, Appl
92	56	16.4	78	US-09-841-132-388	Sequence 388, App

93	56	16.4	229	10	US-09-815-242-5017	Sequence 5017, Ap
94	56	16.4	229	10	US-09-815-242-10613	Sequence 10613, A
95	56	16.4	279	9	US-10-084-205-50	Sequence 50, Appl
96	56	16.4	279	10	US-09-815-242-12136	Sequence 12136, A
97	56	16.4	279	10	US-09-925-637-50	Sequence 50, Appl
98	56	16.4	343	10	US-09-815-242-5705	Sequence 5705, Ap
99	56	16.4	354	10	US-09-815-242-12464	Sequence 12464, A
100	55.5	16.2	133	10	US-09-797-908-8	Sequence 8, Appl1

ALIGNMENTS

RESULT 1
US-09-987-485-2
; Sequence 2, Application US/09987485
; Patent No. US20020142355A1
; GENERAL INFORMATION:
; APPLICANT: Parrot, Michael
; TITLE OF INVENTION: Methods for the In Vivo Biotin Labeling of Polypeptides
; FILE REFERENCE: 15987/268653
; CURRENT APPLICATION NUMBER: US/09/987,485
; CURRENT FILING DATE: 2001-11-14
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 70
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-987-485-2

Query Match 100.0%; Score 342; DB 10; Length 70;
Best Local Similarity 100.0%; Pred. No. 1, 1e-33;
Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EGEIPAPLAGTYSKILVREGDTVKAGQTVLLEANKMETEINAPDGVKVEKYLVERDAV 60
DB 1 EGEIPAPLAGTYSKILVREGDTVKAGQTVLLEANKMETEINAPDGVKVEKYLVERDAV 60
OY 61 OGGGGLIKIG 70
DB 61 OGGGGLIKIG 70

RESULT 2
US-09-987-485-1
; Sequence 1, Application US/09987485
; Patent No. US20020142355A1
; GENERAL INFORMATION:
; APPLICANT: Parrot, Michael
; TITLE OF INVENTION: Methods for the In Vivo Biotin Labeling of Polypeptides
; FILE REFERENCE: 15987/268653
; CURRENT APPLICATION NUMBER: US/09/987,485
; CURRENT FILING DATE: 2001-11-14
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 122
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-987-485-1

Query Match 100.0%; Score 342; DB 10; Length 122;
Best Local Similarity 100.0%; Pred. No. 2, 2e-33;
Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EGEIPAPLAGTYSKILVREGDTVKAGQTVLLEANKMETEINAPDGVKVEKYLVERDAV 60
DB 53 EGEIPAPLAGTYSKILVREGDTVKAGQTVLLEANKMETEINAPDGVKVEKYLVERDAV 112
OY 61 OGGGGLIKIG 70

DB 113 OGGGGLIKIG 122

RESULT 3
US-09-974-973-2
; Sequence 2, Application US/09974973
; Patent No. US20020177202A1
; GENERAL INFORMATION:
; APPLICANT: Hanke, Paul D.
; TITLE OF INVENTION: Feedback-Resistant Pyruvate Carboxylase Gene from Corynebacter
; FILE REFERENCE: 1533.1230001/MAC/RGM
; CURRENT APPLICATION NUMBER: US/09/974,973
; CURRENT FILING DATE: 2001-10-21
; PRIOR APPLICATION NUMBER: US 60/239,913
; PRIOR FILING DATE: 2000-10-13
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 1157
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-974-973-2

Query Match 46.3%; Score 158.5; DB 9; Length 1157;
Best Local Similarity 46.4%; Pred. No. 2, 3e-10;
Matches 32; Conservative 13; Mismatches 23; Indels 1; Gaps 1;

OY 1 EGEIPAPLAGTYSKILVREGDTVKAGQTVLLEANKMETEINAPDGVKVEKYLVERDAV 60
DB 1089 KGHVAPAPFAGVVT-VTVAEGDEVKAGDAVAITEANKMETATITASVDGKIERVVPAAATKV 1147
OY 61 OGGGGLIKI 69
DB 1148 EGGDLIVV 1156

RESULT 4
US-09-974-973-4
; Sequence 4, Application US/09974973
; Patent No. US20020177202A1
; GENERAL INFORMATION:
; APPLICANT: Hanke, Paul D.
; TITLE OF INVENTION: Feedback-Resistant Pyruvate Carboxylase Gene from Corynebacter
; FILE REFERENCE: 1533.1230001/MAC/RGM
; CURRENT APPLICATION NUMBER: US/09/974,973
; CURRENT FILING DATE: 2001-10-21
; PRIOR APPLICATION NUMBER: US 60/239,913
; PRIOR FILING DATE: 2000-10-13
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 1157
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-974-973-4

Query Match 46.3%; Score 158.5; DB 9; Length 1157;
Best Local Similarity 46.4%; Pred. No. 2, 3e-10;
Matches 32; Conservative 13; Mismatches 23; Indels 1; Gaps 1;

OY 1 EGEIPAPLAGTYSKILVREGDTVKAGQTVLLEANKMETEINAPDGVKVEKYLVERDAV 60
DB 1089 KGHVAPAPFAGVVT-VTVAEGDEVKAGDAVAITEANKMETATITASVDGKIERVVPAAATKV 1147
OY 61 OGGGGLIKI 69
DB 1148 EGGDLIVV 1156

RESULT 5
US-09-974-973-19
; Sequence 19, Application US/09974973

Db 126 DGLITIG 132

RESULT 9

US-09-815-242-12361

Sequence 12361, Application US/09815242

Patent No. US20020061569A1

GENERAL INFORMATION:

APPLICANT: Haselbeck, Robert

APPLICANT: Ohlsen, Karl L.

APPLICANT: Zyskind, Judith W.

APPLICANT: Wall, Daniel

APPLICANT: Trawick, John D.

APPLICANT: Cair, Grant J.

APPLICANT: Yamamoto, Robert T.

APPLICANT: Xu, H. Howard

TITLE OF INVENTION: Identification of Essential Genes in

FILE REFERENCE: ELITRA.011A

CURRENT APPLICATION NUMBER: US/09/815,242

PRIOR FILING DATE: 2001-03-21

PRIOR APPLICATION NUMBER: 60/191,078

PRIOR FILING DATE: 2000-03-21

PRIOR APPLICATION NUMBER: 60/206,848

PRIOR FILING DATE: 2000-05-23

PRIOR APPLICATION NUMBER: 60/207,727

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: 60/242,578

PRIOR FILING DATE: 2000-10-23

PRIOR APPLICATION NUMBER: 60/253,625

PRIOR FILING DATE: 2000-11-27

PRIOR APPLICATION NUMBER: 60/257,931

PRIOR FILING DATE: 2000-12-22

PRIOR APPLICATION NUMBER: 60/269,308

NUMBER OF SEQ ID NOS: 14110

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 12361

LENGTH: 1073

TYPE: PRF

ORGANISM: Staphylococcus aureus

US-09-815-242-12361

Query Match

Best Local Similarity 42.7%; Score 146; DB 10; Length 1073;

Matches 31; Conservative 13; Mismatches 22; Indels 0; Gaps 0;

Db 1003 IGAQMPGSTEVEKVSGETVKANQPLITEAMKMETTIOAFDGVIKQVTVNNGDTIATG 1062

OY 64 OGLIKI 69

Db 1063 DLIIEI 1068

RESULT 10

US-09-815-242-5468

Sequence 5468, Application US/09815242

Patent No. US20020061569A1

GENERAL INFORMATION:

APPLICANT: Haselbeck, Robert

APPLICANT: Ohlsen, Karl L.

APPLICANT: Zyskind, Judith W.

APPLICANT: Wall, Daniel

APPLICANT: Trawick, John D.

APPLICANT: Cair, Grant J.

APPLICANT: Yamamoto, Robert T.

APPLICANT: Xu, H. Howard

TITLE OF INVENTION: Identification of Essential Genes in

FILE REFERENCE: ELITRA.011A

CURRENT APPLICATION NUMBER: US/09/815,242

CURRENT FILING DATE: 2001-03-21

PRIOR APPLICATION NUMBER: 60/191,078

PRIOR FILING DATE: 2000-03-21

PRIOR APPLICATION NUMBER: 60/206,848

PRIOR FILING DATE: 2000-05-23

PRIOR APPLICATION NUMBER: 60/207,727

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: 60/242,578

PRIOR FILING DATE: 2000-10-23

PRIOR APPLICATION NUMBER: 60/253,625

PRIOR FILING DATE: 2000-11-27

PRIOR APPLICATION NUMBER: 60/257,931

PRIOR FILING DATE: 2000-12-22

PRIOR APPLICATION NUMBER: 60/269,308

PRIOR FILING DATE: 2001-02-16

NUMBER OF SEQ ID NOS: 14110

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 5468

LENGTH: 1147

TYPE: PRF

ORGANISM: Staphylococcus aureus

US-09-815-242-5468

Query Match

Best Local Similarity 42.7%; Score 146; DB 10; Length 1147;

Matches 31; Conservative 13; Mismatches 22; Indels 0; Gaps 0;

Db 1079 IGAQMPGSTEVEKVSGETVKANQPLITEAMKMETTIOAFDGVIKQVTVNNGDTIATG 1138

OY 64 OGLIKI 69

Db 1139 DLIIEI 1144

RESULT 11

US-10-051-643-166

Sequence 166, Application US/10051643

Publication No. US20020197265A1

GENERAL INFORMATION:

APPLICANT: Watson, James D.

APPLICANT: Tan, Paul L. J.

TITLE OF INVENTION: Methods and Compounds for the Treatment

TITLE OF INVENTION: of Immunologically-Mediated Diseases of the Respiratory

FILE REFERENCE: 11000.1008c2

CURRENT APPLICATION NUMBER: US/10/051,643

PRIOR FILING DATE: 2002-01-18

PRIOR APPLICATION NUMBER: US09/156,181

PRIOR FILING DATE: 1998-09-17

PRIOR APPLICATION NUMBER: US 08/996,624

PRIOR FILING DATE: 1997-12-23

NUMBER OF SEQ ID NOS: 208

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 166

LENGTH: 108

TYPE: PRF

ORGANISM: Mycobacterium vaccae

US-10-051-643-166

Query Match

Best Local Similarity 42.1%; Score 144; DB 9; Length 108;

Matches 32; Conservative 9; Mismatches 26; Indels 0; Gaps 0;

Db 37 VDAPFASVWKVDVAGDRVAGQPLALTEAMKMETVLAAPADGVYQTLVSAGHLVDG 96

OY 64 OGLIKIG 70

Db 97 TPLVVG 103

```
RESULT 12
US-09-880-505-166
; Sequence 166, Application US/09880505.
; Publication No. US20030007976A1
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Tan, Paul L.J.
; APPLICANT: Prestidge, Ross
; TITLE OF INVENTION: Methods and Compounds for the Treatment
; FILE REFERENCE: 11000.1007c2
; CURRENT APPLICATION NUMBER: US/09/880,505
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 09/324,542
; PRIOR FILING DATE: 1999-06-02
; PRIOR APPLICATION NUMBER: US 08/997,080
; PRIOR FILING DATE: 1997-12-23
; NUMBER OF SEQ ID NOS: 194
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 166
; LENGTH: 108
; TYPE: PRF
; ORGANISM: Mycobacterium vaccae
US-09-880-505-166

Query Match          42.1%; Score 144; DB 9; Length 108;
Best Local Similarity 47.8%; Pred. No. 6.2e-10;
Matches 32; Conservative 9; Mismatches 26; Indels 0; Gaps 0;

QY 4 IPAPLACTVSKIVKEDGYKAGQTVLVLEAMKETEINAPTDGKVERKVDAYGG 63
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 37 VDAPEFASVGVYVLAAGPDKPAFVTVGDSVKKGQTVLIEAMKETEINAP 96
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 64 QGLIKIG 70
    ||:|
Db 97 TPLVVG 103
    ||:|

RESULT 13
US-09-815-242-13363
; Sequence 1363, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
```

```
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13363
; LENGTH: 161
; TYPE: PRF
; ORGANISM: Streptococcus pneumoniae
US-09-815-242-13363

Query Match          42.1%; Score 144; DB 10; Length 161;
Best Local Similarity 41.6%; Pred. No. 1e-09;
Matches 32; Conservative 15; Mismatches 22; Indels 8; Gaps 2;

QY 1 EGE-IPAPLACTV-----SKIVKEDGYKAGQTVLVLEAMKETEINAPTDGK 52
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 84 EGNLVESPLVGVYVLAAGPDKPAFVTVGDSVKKGQTVLIEAMKETEINAP 143
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 53 LYKRDVAVGGGGLIKI 69
    ||:|
Db 144 LVSNEMVERGKGLVRI 160
    ||:|

RESULT 14
US-09-815-242-13615
; Sequence 13615, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13615
; LENGTH: 161
; TYPE: PRF
; ORGANISM: Streptococcus pneumoniae
US-09-815-242-13615

Query Match          42.1%; Score 144; DB 10; Length 161;
Best Local Similarity 41.6%; Pred. No. 1e-09;
Matches 32; Conservative 15; Mismatches 22; Indels 8; Gaps 2;

QY 1 EGE-IPAPLACTV-----SKIVKEDGYKAGQTVLVLEAMKETEINAPTDGK 52
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 84 EGNLVESPLVGVYVLAAGPDKPAFVTVGDSVKKGQTVLIEAMKETEINAP 143
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 53 LYKRDVAVGGGGLIKI 69
    ||:|
Db 144 LVSNEMVERGKGLVRI 160
    ||:|
```

RESULT 15
US-10-051-643-199
; Sequence 199, Application US/10051643
; Publication No. US20020197265A1
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Tan, Paul L. J.
; TITLE OF INVENTION: Methods and Compounds for the Treatment
; TITLE OF INVENTION: of Immunologically-Mediated Diseases of the Respiratory
; FILE REFERENCE: 11000.1008c2
; CURRENT APPLICATION NUMBER: US/10/051,643
; CURRENT FILING DATE: 2002-01-18
; PRIOR APPLICATION NUMBER: US09/156,181
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: US 08/996,624
; PRIOR FILING DATE: 1997-12-23
; NUMBER OF SEQ ID NOS: 208
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 199
; LENGTH: 243
; TYPE: PRF
; ORGANISM: Mycobacterium vaccae
US-10-051-643-199

Query Match 42.1%; Score 144; DB 9; Length 243;
Best Local Similarity 47.8%; Pred. No. 1.7e-09;
Matches 32; Conservative 9; Mismatches 26; Indels 0; Gaps 0;

QY 4 IIPPLAGTVSKILVKGDPYKAGOTVLYEAMKMETEINAPDQKVEKYVERDVAOG 63
DB 172 VDAPFSSVWKVYVAVGDRVAGOPPLALEAMKMETVLRAPAGVVTQIIVSGHLVDPG 231
QY 64 OGILIKIG 70
DB 232 TPLVYVG 238

RESULT 16
US-09-815-242-10806
; Sequence 10806, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10806

LENGTH: 1142
; TYPE: PRF
; ORGANISM: Enterococcus faecalis
US-09-815-242-10806

Query Match 42.1%; Score 144; DB 10; Length 1142;
Best Local Similarity 43.3%; Pred. No. 1.2e-08;
Matches 29; Conservative 17; Mismatches 21; Indels 0; Gaps 0;

QY 3 EIPAPLAGTVSKILVKGDPYKAGOTVLYEAMKMETEINAPDQKVEKYVERDVAOG 62
DB 1073 QIGATMSGSTQVLYVRGDKVEKGPPLITTEAMKMETTEARRAGTVDHIVYEGEAISS 1132
QY 63 GGGILIKI 69
DB 1133 GDLLLEV 1139

RESULT 17
US-10-160-501-17
; Sequence 17, Application US/10160501
; Publication No. US20030059919A1
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Rudolph-Owen, Laura A.
; APPLICANT: Kapeller-Libermann, Rosana
; TITLE OF INVENTION: NOVEL HUMAN 39228, 21956, 25856, 22244, 8701, 32263,
; TITLE OF INVENTION: 50250, 55158, 47765, 62088, 50566, AND 48118
; FILE REFERENCE: MNT-250
; CURRENT APPLICATION NUMBER: US/10/160,501
; CURRENT FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: US 09/838,573
; PRIOR FILING DATE: 2002-04-18
; PRIOR APPLICATION NUMBER: US 60/197,747
; PRIOR FILING DATE: 2000-04-18
; PRIOR APPLICATION NUMBER: US 09/870,133
; PRIOR FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: US 60/207,649
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/870,130
; PRIOR FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: US 60/207,640
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/862,535
; PRIOR FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: US 60/205,961
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: US 09/870,383
; PRIOR FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: US 60/207,506
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/860,821
; PRIOR FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: US 60/205,449
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: US 09/870,110
; PRIOR FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: US 60/207,650
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/907,509
; PRIOR FILING DATE: 2001-07-16
; PRIOR APPLICATION NUMBER: US 60/218,385
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/945,327
; PRIOR FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: US 60/229,425
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: US 60/318,581
; PRIOR FILING DATE: 2001-09-10
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: FastSeq Version 4.0
; SEQ ID NO 17

LENGTH: 725
TYPE: PRT
ORGANISM: Homo sapiens
US-10-160-501-17

Query Match 41.8%; Score 143; DB 9; Length 725;
Best Local Similarity 52.9%; Pred. No. 8.9e-09;
Matches 27; Conservative 10; Mismatches 14; Indels 0; Gaps 0;

OY 6 APLAGTIVKEDPTVAKAGQVLEAMKMEINAPPTGKVEKLVK 56
DB 651 APMGTIEKVFVKAGDKVAKGDSLMVIMAMKMEITIKSPDKGVKKVYRE 701

RESULT 18
US-09-815-242-4962
Sequence 4962, Application US/09815242
Patent No. US20020061569A1

GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
Prokaryotes
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 4962
LENGTH: 160
TYPE: PRT
ORGANISM: Enterococcus faecalis
US-09-815-242-4962

Query Match 40.8%; Score 139.5; DB 10; Length 160;
Best Local Similarity 40.5%; Pred. No. 3.5e-09;
Matches 30; Conservative 12; Mismatches 25; Indels 7; Gaps 1;

OY 3 EIPPLAGTV-----SKILVREGDPYKAGQVLEAMKMEINAPPTGKVEKLVK 55
DB 87 EITSPIVIGVILQAPDPKRENVKVGDIYKTDVCIYVAMKLMNEITATVDGVITTEILVN 146

OY 56 ERDAVGGGGLIKI 69
DB 147 NEDVVEFGQPLFRV 160

RESULT 19
US-09-815-242-10922
Sequence 10922, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert

APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
Prokaryotes
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 10922
LENGTH: 162
TYPE: PRT
ORGANISM: Enterococcus faecalis
US-09-815-242-10922

Query Match 40.8%; Score 139.5; DB 10; Length 162;
Best Local Similarity 40.5%; Pred. No. 3.6e-09;
Matches 30; Conservative 12; Mismatches 25; Indels 7; Gaps 1;

OY 3 EIPPLAGTV-----SKILVREGDPYKAGQVLEAMKMEINAPPTGKVEKLVK 55
DB 85 EITSPIVIGVILQAPDPKRENVKVGDIYKTDVCIYVAMKLMNEITATVDGVITTEILVN 144

OY 56 ERDAVGGGGLIKI 69
DB 145 NEDVVEFGQPLFRV 158

RESULT 20
US-09-767-479-2
Sequence 2, Application US/09767479
Patent No. US20010036654A1
GENERAL INFORMATION:
APPLICANT: Haselkorn, Robert
Gornicki, Piotr

TITLE OF INVENTION: ACETYL-CoA CARBOXYLASE COMPOSITIONS AND
METHODS OF USE

NUMBER OF SEQUENCES: 29

CORRESPONDENCE ADDRESS:

ADDRESSEE: Arnold, White & Durkee

STREET: P.O. Box 4433

CITY: Houston

STATE: Texas

COUNTRY: United States of America

ZIP: 77210

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/767,479

FILING DATE: 22-Jan-2001

CLASSIFICATION: Unknown

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/468,793
FILING DATE: <UNKNOWN>
APPLICATION NUMBER: US SN 07/956,700
FILING DATE: 02-OCT-1992
APPLICATION NUMBER: PCT/US93/09340
FILING DATE: 30-SEP-1993
ATTORNEY/AGENT INFORMATION:
NAME: Kitchell, Barbara S.
REGISTRATION NUMBER: 33,928
REFERENCE/DOCKET NUMBER: ARCD:152/KIT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (713) 789-2679
TELEX: 79-0924
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 182 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-767-479-2

Query Match 37.0%; Score 126.5; DB 10; Length 182;
Best Local Similarity 35.1%; Pred. No. 1.4e-07;
Matches 26; Conservative 19; Mismatches 22; Indels 7; Gaps 1;

QY 3 EIRAPLAGTVSK-----ILVKEGDTVRKAGQTVLVLEAMKMETEINAPTDGKVEVLYK 55
DB 106 EVASPMVGTFYRAPAPGAEVFEVGRIRGQTVCTIEAKKLNLEADVSGVITLVO 165
QY 56 ERDAVGGGGLIKI 69
DB 166 NGEPVEYNQPLMRI 179

RESULT 21
US-09-738-626-6940
Sequence 6940, Application US/09738626
Publication No. US20020197605A1
GENERAL INFORMATION:
APPLICANT: NAKAGAWA, SATOSHI
APPLICANT: MIZOGUCHI, HIROSHI
APPLICANT: ANDO, SEIKO
APPLICANT: HAYASHI, MIKIRO
APPLICANT: OCHIAI, KEIKO
APPLICANT: YOKOI, HARUHIKO
APPLICANT: TATEISHI, NAKO
APPLICANT: SENOH, AKIHIRO
APPLICANT: IKEDA, MASATO
FILE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-125
CURRENT APPLICATION NUMBER: US/09/738,626
PRIOR FILING DATE: 2000-12-18
PRIOR APPLICATION NUMBER: JP 99/377484
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: JP 00/159162
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: JP 00/280988
PRIOR FILING DATE: 2000-08-03
NUMBER OF SEQ ID NOS: 7059
SOFTWARE: PatentIn ver. 3.0
SEQ ID NO 6940
LENGTH: 591
TYPE: PRT
ORGANISM: Corynebacterium glutamicum
US-09-738-626-6940

Query Match 36.8%; Score 126; DB 9; Length 591;
Best Local Similarity 43.9%; Pred. No. 7.2e-07;

Matches 29; Conservative 10; Mismatches 27; Indels 0; Gaps 0;
QY 4 IPAPLAGTVSKILVKEGDTVRKAGQTVLVLEAMKMETEINAPTDGKVEVLYK 63
DB 525 VAAPMGTVYKVAVEGAEVNEGDTVYVLEAMKMEPVNAHSGVTGLTVAAGEVNG 584
QY 64 OGGLIKI 69
DB 585 VVLEI 590

RESULT 22
US-09-815-242-12581
Sequence 12581, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE OF INVENTION: Prokaryotes
CURRENT APPLICATION NUMBER: US/09/815,242
PRIOR FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 12581
LENGTH: 154
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-09-815-242-12581

Query Match 34.9%; Score 119.5; DB 10; Length 154;
Best Local Similarity 39.7%; Pred. No. 7.9e-07;
Matches 29; Conservative 12; Mismatches 25; Indels 7; Gaps 1;

QY 4 IPAPLAGTVSK-----ILVKEGDTVRKAGQTVLVLEAMKMETEINAPTDGKVEVLYK 56
DB 81 INAPMGTVYKSPSPDEAVYVGVDSNETTVCTIEAKKLNLEADVSGVITLVO 140
QY 57 RDAVGGGGLIKI 69
DB 141 GOMETGQPLFRV 153

RESULT 23
US-09-767-479-4
Sequence 4, Application US/09767479
Patent No. US2001003654A1
GENERAL INFORMATION:
APPLICANT: Haselkorn, Robert
Gornicki, Piotr
TITLE OF INVENTION: ACETYL-CoA CARBOXYLASE COMPOSITIONS AND
METHODS OF USE

NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: United States of America
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/767,479
FILING DATE: 22-Jan-2001
CLASSIFICATION: Unknown
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/468,793
FILING DATE: <Unknown>
APPLICATION NUMBER: US SN 07/956,700
FILING DATE: 02-OCT-1992
APPLICATION NUMBER: PCT/US93/09340
FILING DATE: 30-SEP-1993
ATTORNEY/AGENT INFORMATION:
NAME: Kitchell, Barbara S.
REGISTRATION NUMBER: 33,928
REFERENCE/DOCKET NUMBER: ARCD:152/KIT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (713) 789-2679
TELEX: 79-0924
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 158 amino acids
TYPE: amino acid
STRANDEDNESS: single
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-767-479-4
Query Match 34.9%; Score 119.5; DB 10; Length 158;
Best Local Similarity 35.1%; Pred. No. 8.2e-07;
Matches 26; Conservative 16; Mismatches 25; Indels 7; Gaps 1;
QY 3 EIPAPLAGTSKTI-----LYKGGDTYKAGQTVLVEAMKMEIEINAPDGVKYEKVLK 55
DB 82 EIRAPMGITYYRAPAPEPPVNVGDIQVGQVCIIEAMKMEIESEVTEGVEVILVO 141
QY 56 ERDAVGGGLIKI 69
DB 142 NGEVEFNGPLFRL 155
RESULT 24
US-09-791-171-48
Sequence 48, Application US/09791171
Patent No. US20020094336A1
GENERAL INFORMATION:
APPLICANT: ANDERSEN, Peter
APPLICANT: NIELSEN, Rikke
APPLICANT: OETTINGER, Thomas
APPLICANT: RASMUSSEN, Peter Birk
APPLICANT: ROSENKRANDS, Ida
APPLICANT: WELDINGH, Karin
APPLICANT: FIORIO, Walter
TITLE OF INVENTION: NUCLEIC ACIDS FRAGMENTS AND POLYPEPTIDE FRAGMENTS
FILE REFERENCE: 670001-2002.1
CURRENT APPLICATION NUMBER: US/09/791,171
CURRENT FILING DATE: 2001-02-20
PRIOR APPLICATION NUMBER: 09/050,739

PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 0376/97
PRIOR FILING DATE: 1997-04-02
PRIOR APPLICATION NUMBER: 1277/97
PRIOR FILING DATE: 1997-11-10
PRIOR APPLICATION NUMBER: 60/044,624
PRIOR FILING DATE: 1997-04-18
PRIOR APPLICATION NUMBER: 60/070,488
PRIOR FILING DATE: 1998-01-05
NUMBER OF SEQ ID NOS: 173
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 48
LENGTH: 71
TYPE: PRT
ORGANISM: Mycobacterium tuberculosis
US-09-791-171-48
Query Match 33.6%; Score 115; DB 10; Length 71;
Best Local Similarity 37.7%; Pred. No. 1e-06;
Matches 23; Conservative 13; Mismatches 25; Indels 0; Gaps 0;
QY 3 EIPAPLAGTSKTIKVGDTYKAGQTVLVEAMKMEIEINAPDGVKYEKVLK 62
DB 4 DVREIYASVLEVVNEDGQDIDKGDVYVLLSKMEIPIVLAEAAGTVSKVAVSVDVIA 63
QY 63 G 63
DB 64 G 64
RESULT 25
US-09-815-242-10036
Sequence 10036, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 10036
LENGTH: 630
TYPE: PRT
ORGANISM: Escherichia coli
US-09-815-242-10036
Query Match 33.2%; Score 113.5; DB 10; Length 630;
Best Local Similarity 34.0%; Pred. No. 2.4e-05;
Matches 32; Conservative 9; Mismatches 26; Indels 27; Gaps 3;

Best Local Similarity 31.9%; Pred. No. 0.00036;
Matches 22; Conservative 15; Mismatches 27; Indels 5; Gaps 1.

```

QY      1  EGELPAPLAGVSKILYKEDGTFYKAGQTVLVLEAMKMETEINAPTDKVEKVLVERDAV 600
      |  :  |  :  |  |  |  |  |  :  |  :  |  |  |  |  |  |  |  |  :  :
Db      14  ESILEAPI-----RVSVKIGDSIKQGVLFIIETDRTSLEIVSPVDGTVSVKFIADDEII 688

```

```

QY      61 QGGGGLIKI 69
      : | | |
Db      69 ERDQLCTI 77

```

RESULT 29
US-09-738-626-5917

```

1  APPLICANT: NAKAGAMA, SATOSHI
2  APPLICANT: MIZOGUCHI, HIROSHI
3  APPLICANT: ANDO, SEIKO
4  APPLICANT: HATAISHI, MIKIRO
5  APPLICANT: OCHIAI, KEIKO
6  APPLICANT: YOKOI, HARUHIKO
7  APPLICANT: TATEISHI, NAOKO
8  APPLICANT: SENOH, AKIHIRO
9  APPLICANT: IKEDA, MASATO
10 APPLICANT: OZAWA, AKIO
11 TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
12 FILE REFERENCE: 248-125
13 CURRENT APPLICATION NUMBER: US/09/738, 626
14 CURRENT FILING DATE: 2000-12-18
15 PRIOR APPLICATION NUMBER: JP 99/377484
16 PRIOR FILING DATE: 1999-12-16
17 PRIOR APPLICATION NUMBER: JP 00/159162
18 PRIOR FILING DATE: 2000-04-07
19 PRIOR APPLICATION NUMBER: JP 00/280988
20 PRIOR FILING DATE: 2000-08-03
21 NUMBER OF SEQ ID NOS: 7059
22 SOFTWARE: PatentIn ver. 3.0
23 SEQ ID NO 5917
24 LENGTH: 675
25 TYPE: PRT
26 ORGANISM: Corynebacterium glutamicum
27
28 US-09-738-626-5917

```

Query Match	28.9%	Score 99	DB 9	Length 675
Best Local Similarity	34.4%	Pred. No. 0.0014		
Matches	21	Conservative	15	Mismatches 25; Indels 0; Gaps 0

Qy 10 GTVSKILVKEGDTYKAGQTVLVLEAMKETEINLPDTGKVEKLVKERDAVGGGGLIKI 69
 ||:::| |||: : : : |:||| : | : : : | : : : |
 Db 252 GTIQWLKAVGDTVEYDEPLLEVSTDKVDTEIPVAGTIVEILWDEDDTVGVGAVIARI 311

QY	70	G	70
		1	
Db	312	G	312

RESULT 30
US-09-815-242-12062
Sequence 12062, Application US/09815242

1 APPLICANT: Haselbeck, Robert
 2 APPLICANT: Ohlsen, Karl L.
 3 APPLICANT: Zykkind, Judith W.
 4 APPLICANT: Wall, Daniel
 5 APPLICANT: Trawick, John D.
 6 APPLICANT: Carr, Grant J.
 7 APPLICANT: Yamamoto, Robert T.
 8 APPLICANT: Xu, H. Howard
 9
 10 TITLE OF INVENTION: Identification of Essential Genes in
 11 TITLE OF INVENTION: Prokaryotes

```

; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-01

```

Query Match	28.88;	Score 98.5;	DB 10;	Length 156;
Best Local Similarity	32.98;	Pred. No. 0.00025;		
Matches	24;	Conservative	12;	Mismatches 30;
				Indels 7;
				Gaps 1

```

QY  4 IPAPLAG-----VSILVKEEDFYACGTVLALMKMETIINPPTGKVEKVLAK 56
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db  83 VRSPWGTGYRAASPTSANFYVGQSVKKDILCLIVEAMKMNHHIEAVSSTIESLVEN 142
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY  57 RDAVGGGLIKI 69
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db  143 GQVVEFDQPLETI 155
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

Search completed: May 1, 2003, 08:16:11
Job time : 135.531 secs

GenCore version 5.1.4-p5-4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 1, 2003, 07:49:48 ; Search time 66.7188 Seconds
(without alignments)
376.772 Million cell updates/sec

Title: US-09-987-485-1

Perfect score: 615

Sequence: 1 MKIKVTNGTAYDVVDVVK.....KVLKERDAVGGGGLIKIG 122

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 segs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database :

SPREMBL_21:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriophage:*
17: sp_archaeop:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	231.5	37.6	145	17	Q9V0A6
2	221	35.9	149	17	Q59021
3	218	35.4	144	17	Q8U303
4	205	33.3	129	2	Q57111
5	198.5	32.3	134	2	Q54030
6	196.5	32.0	145	2	Q9ZAA7
7	190.5	31.0	655	16	Q67484
8	190	30.9	140	17	Q28067
9	188	30.6	116	16	Q99ZL6
10	187.5	30.5	599	16	Q9P000
11	185	30.1	597	16	Q9K0H1
12	184	29.9	571	17	Q58564
13	179.5	29.2	122	16	Q8R7M0
14	176.5	28.7	576	16	Q8U917
15	176	28.6	134	16	Q9WZB6
16	176	28.6	596	2	Q48826

17	176	28.6	984	2	Q9XB11
18	175.5	28.5	610	17	Q9HPP8
19	174	28.3	591	17	Q8R0N4
20	173.5	28.2	132	16	Q99ZL1
21	173.5	28.2	186	17	Q97YV7
22	173.5	28.2	591	16	Q8XCX8
23	172	28.0	134	16	Q8R3Y8
24	170.5	27.7	591	2	P71122
25	170	27.6	1144	16	Q97ER7
26	168.5	27.4	607	16	Q9HDL1
27	168.5	27.4	620	16	Q67544
28	168	27.3	135	16	Q8RAJ2
29	168	27.3	158	16	Q920K0
30	166	27.0	600	16	P96890
31	165.5	26.9	598	2	Q8R0N9
32	165	26.8	1150	16	Q9K9M0
33	164	26.7	654	16	Q9AC65
34	163	26.5	142	17	Q28194
35	162	26.3	169	17	Q974R8
36	162	26.3	597	2	Q54105
37	162	26.3	597	2	Q8R0L2
38	162	26.3	1139	2	Q8R0X3
39	161.5	26.3	1174	16	Q8UBX3
40	159.5	25.9	1140	2	Q54587
41	159	25.9	665	16	Q92H13
42	158.5	25.8	593	16	Q83095
43	157	25.5	161	16	Q9FBC1
44	157	25.5	170	16	Q9A743
45	157	25.5	1146	16	Q92CWL
46	157	25.5	1146	16	Q8Y646
47	156	25.4	1150	16	Q990Y8
48	155	25.2	167	1	Q52603
49	154.5	25.1	1124	16	Q9RK64
50	154.5	25.1	1152	16	Q98R27
51	153.5	25.0	436	10	Q40121
52	152.5	24.8	162	16	Q8URF6
53	152.5	24.8	602	16	Q9CL25
54	152	24.7	655	16	Q91299
55	151.5	24.6	1192	3	Q9HE88
56	151	24.6	166	16	Q99YD8
57	150	24.4	70	16	Q9KPS8
58	150	24.4	155	16	Q9CHP5
59	150	24.4	1154	2	Q59740
60	150	24.4	1154	2	Q59740
61	149	24.2	1147	2	P94448
62	148	24.1	1152	16	Q92L13
63	147	23.9	694	16	Q985D4
64	147	23.9	1078	16	Q9A3T0
65	146.5	23.8	712	16	Q8Y2S0
66	146.5	23.8	1127	16	P95127
67	145.5	23.7	1193	3	Q93918
68	145	23.6	156	16	Q8YCU5
69	144.5	23.5	1137	2	Q9RAT6
70	144	23.4	661	16	Q9HZV6
71	144	23.4	1185	3	P78822
72	144	23.4	1185	3	Q9U0E1
73	143.5	23.3	1137	2	Q9F843
74	143.5	23.3	1137	16	Q9CH07
75	142.5	23.2	572	17	Q8TSL1
76	141.5	23.0	678	16	Q8YEF3
77	141	22.9	601	2	Q9RF99
78	140	22.8	1175	3	Q8X1R3
79	139.5	22.7	155	16	Q8XVP4
80	139	22.6	187	16	Q9RV33
81	136.5	22.2	158	2	Q54761
82	134.5	21.9	456	2	Q9LB26
83	134.5	21.9	1181	5	Q9XZ00
84	134.5	21.9	1196	5	Q9Y7E9
85	134	21.8	1180	13	Q9DDT1
86	133.5	21.7	159	16	Q97DA8
87	133	21.6	920	2	Q9KWU5
88	133	21.6	1148	16	Q9KWU4
89	132	21.5	152	16	Q8KRB7

Q9XB11	bacillus ce
Q9HPP8	halobacteri
Q8R0N4	corynebacte
Q99ZL1	streptococ
Q97YV7	sulfolobus
Q8XCX8	salmonella
Q8R3Y8	fusobacteri
P71122	corynebacte
Q97ER7	clostridium
Q9HDL1	pseudomonas
Q67544	aquifex aeo
Q8RAJ2	thermoanaer
Q920K0	rhizobium m
P96890	mycobacteri
Q8R0N9	mycobacteri
Q9K9M0	bacillus ha
Q9AC65	caulobacter
Q28194	archaeoglob
Q974R8	sulfolobus
Q54105	saccharopol
Q8R0L2	corynebacte
Q8UBX3	agrobacteri
Q54587	corynebacte
Q92H13	ricketsia
Q83095	treponema p
Q9FBC1	streptococ
Q9A743	caulobacter
Q92CWL	listeria in
Q8Y646	listeria mo
Q990Y8	staphylococ
Q52603	sulfolobus
Q9RK64	streptomyce
Q98R27	rhizobium l
Q40121	lycopersico
Q8URF6	agrobacteri
Q9CL25	pasteurella
Q91299	pseudomonas
Q9HE88	aspergillus
Q99YD8	streptococ
Q9KPS8	bacillus ha
Q9CHP5	lactococcus
Q9ZCU3	ricketsia
Q59740	rhizobium e
P94448	bacillus st
Q92L13	rhizobium m
Q985D4	rhizobium l
Q9A3T0	caulobacter
Q8Y2S0	raistonia s
P95127	mycobacteri
Q93918	aspergillus
Q8YCU5	brucella me
Q9RAT6	lactococcus
Q9HZV6	pseudomonas
P78822	schizosacch
Q9U0E1	mycobacteri
Q9F843	lactococcus
Q9CH07	methanosarc
Q8TSL1	brucella me
Q9RF99	haemophilus
Q8X1R3	pichia angu
Q8XVP4	raistonia s
Q9RV33	deinococcus
Q54761	synecococc
Q9LB26	rhodospirill
Q9XZ00	drosophila
Q9Y7E9	brachydanio
Q9DDT1	clostridium
Q97DA8	clostridium
Q9KWU5	bacillus su
Q9KWU4	bacillus su
Q8KRB7	raistonia s

90	131	21.3	573	1	Q9HH18
91	131	21.3	1207	16	Q9A797
92	130	21.1	203	10	Q8W2G0
93	130	21.1	665	5	Q60966
94	130	21.1	1158	16	Q8YJ20
95	129	21.0	198	2	Q9AGU8
96	129	21.0	667	16	P71538
97	129	21.0	935	11	Q62043
98	128.5	20.9	1124	2	Q50450
99	128	20.8	688	5	Q45430
100	127	20.7	590	2	Q9RG06

ALIGNMENTS

RESULT 1

09Y0A6 PRELIMINARY; PRT; 145 AA.

AC 09Y0A6; 01-MAY-2000 (TREMblrel. 13, Created)

DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)

DE 01-MAR-2002 (TREMblrel. 20, Last annotation update)

GN Methylnalonyl-CoA decarboxylase gamma chain.

OS Pyrococcus abyssi.

OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;

OC Pyrococcus.

OX NCBI_TaxID=29292;

RP SEQUENCE FROM N.A.

RC STRAIN=ORSAY;

RA Helling R.;

RT "Pyrococcus abyssi genome sequence: Insights into archaeal chromosome structure and evolution."

RL Submitted (JUL-1999) to the EMBL/Genbank/DBJ databases.

DR EMBL: AJ248285; CAB49799.1; -

DR HSSP: P02905; IBD0.

DR InterPro: IPR001882; Biotin_attach.

DR InterPro: IPR000089; Biotin_lipoyl.

DR Pfam: PF00364; biotin_lipoyl.1.

DR PROSITE: PS00188; BIOTIN; UNKNOWN_1.

KW Complete proteome.

SQ SEQUENCE 145 AA; 15489 MW; 9C1443363F40D94 CRC64;

Query Match 37.6%; Score 231.5; DB 17; Length 145;

Best Local Similarity 41.4%; Pred. No. 2.4e-11;

Matches 60; Conservative 18; Mismatches 44; Indels 23; Gaps 3;

QY 1 MKLKTVNGTAVDVIDV-----VDKSHENPMGTLIFGGGTGGA 40

DB 1 MKKVVVNGKEVEVEEVPMPGKFRVTLGEGYEVETSGAGFYSPROVOVPPAPPPA 60

OY 41 PA---AGAGAGAGAGGEGEIPAPLAGTAVSKILVKEGDTVAGQTVLVLEAMKMETEINAP 97

DB 61 PTPVQAPPTTPVOVASENVVTAPEPKVTLVVOEGQVKLGGLTLLEAMKMETEINAP 120

OY 98 DSKVEKVLVKEKDAVGGGGLIKIG 122

DB 121 DGVVKRIIVKEGDAVDTGTPLEIG 145

RESULT 2

059021 PRELIMINARY; PRT; 149 AA.

AC 059021; 01-AUG-1998 (TREMblrel. 07, Created)

DT 01-AUG-1998 (TREMblrel. 07, Last sequence update)

DE 01-MAR-2002 (TREMblrel. 20, Last annotation update)

GN 149PA long hypothetical methylmalonyl-CoA decarboxylase gamma chain.

OS Pyrococcus horikoshii.

OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;

OC Pyrococcus.

OX NCBI_TaxID=53953;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=OT3;

RX MEDLINE=98344137; PubMed=9679194;

RA Kawarabayashi Y., Sawada M., Horikawa H., Hino Y.,

RA Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai Y.,

RA Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohfuku Y.,

RA Funahashi T., Tanaka T., Kudoh Y., Yamazaki J., Kushida N., Oguchi A.,

RA Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,

RA Masuchi Y., Shizuya H., Kikuchi H.,

RT "Complete sequence and gene organization of the genome of a hyper-thermophilic archaeobacterium, Pyrococcus horikoshii OT3."

RL DNA Res. 5:55-76(1998).

DR EMBL: AP000005; BAA30387.1; -

DR HSSP: P10802; IIVT.

DR InterPro: IPR001882; Biotin_attach.

DR InterPro: IPR000089; Biotin_lipoyl.

DR Pfam: PF00364; biotin_lipoyl.1.

DR PROSITE: PS00188; BIOTIN; UNKNOWN_1.

KW Complete proteome.

SQ SEQUENCE 149 AA; 15985 MW; 1C3A5F47E6BA6F1 CRC64;

Query Match 35.9%; Score 221; DB 17; Length 149;

Best Local Similarity 36.3%; Pred. No. 1.7e-10;

Matches 53; Conservative 21; Mismatches 48; Indels 24; Gaps 1;

QY 1 MKLKTVNGTAVDVIDV-----VDKSHENPMGTLIFGGGTG 36

DB 4 MKKVVVNGKEVEVEEVPMPGKFRVTLGEGYEVETSGAGFYSPROVOVPPAPPPA 63

OY 37 GAPAPAGAGAGAGGEGEIPAPLAGTAVSKILVKEGDTVAGQTVLVLEAMKMETEINAP 96

DB 64 PPTPAPAPSSKTVSENVVSAPEPKVTLVTVGDDRVVGGGLVLEAMKMETEINAP 123

OY 97 TDKVEKVLVKEKDAVGGGGLIKIG 122

DB 124 RDGVVKRIIVKEGDAVDTGTPLEIG 149

RESULT 3

080303 PRELIMINARY; PRT; 144 AA.

AC 080303; 01-JUN-2002 (TREMblrel. 21, Created)

DT 01-JUN-2002 (TREMblrel. 21, Last sequence update)

DE 01-JUN-2002 (TREMblrel. 21, Last annotation update)

GN Methylnalonyl-CoA decarboxylase gamma chain.

OS Pyrococcus furiosus.

OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;

OC Pyrococcus.

OX NCBI_TaxID=2261;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=VCI / DSM 3638 / ATCC 43587 / JCM 8422;

RA Weiss R.B., Dunn D.M., Robb F.T., Brown J.R.;

RT "The complete sequence of the Pyrococcus furiosus genome."

RL Submitted (FEB-2002) to the EMBL/Genbank/DBJ databases.

DR EMBL: AE010188; AAL80797.1; -

KW Complete proteome.

SQ SEQUENCE 144 AA; 15315 MW; 422C96A8ED809C6A CRC64;

Query Match 35.4%; Score 218; DB 17; Length 144;

Best Local Similarity 36.7%; Pred. No. 2.9e-10;

Matches 54; Conservative 21; Mismatches 44; Indels 28; Gaps 3;

QY 1 MKLKTVNGTAVDVIDV-----VDKSHENPMGTLIFGGGTG 38

DB 1 MKKVIINKEVEVEEELMPGKFRVTLGEGYEVETSGAGFYSPROVOVPPAPPPA 57

OY 39 PAPA---AGAGAGAGAGGEGEIPAPLAGTAVSKILVKEGDTVAGQTVLVLEAMKMETEINAP 95

DB 58 PAPAPMTSTPASPQVGNVNTAPKPKILTLVSEGRVITGGLLILEMKMENEIPS 117
QY 96 PTDGKVEKVLKERDAVGGGLIKIRG 122
DB 118 PKDGVKKIYKEGDVTGTGGLIFELG 144

RESULT 4
057111 PRELIMINARY; PRT; 129 AA.
AC 057111:
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)
DE Gamma-subunit, methylmalonyl-CoA decarboxylase.
OS Vellionella parvula.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC Clostridiales; Acidaminococcaceae; Vellionella.
OX NCBI_TaxID=29466;
RN [1]
RP SEQUENCE FROM N.A.
RL J. Biol. Chem. 0:0-0(1993).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=94043308; PubMed=82271015;
RA Huder J.B., Dimroth P.;
RT "Sequence of the sodium ion pump methylmalonyl-CoA decarboxylase from
RT Vellionella parvula."
RL J. Biol. Chem. 268:24564-24571(1993).
CC -I- COFACTOR: BIOTIN (BY SIMILARITY).
DR EMBL; Z24754; CAAB0875.1; -
DR EMBL; L22208; AAC36823.1; -
DR HSSP; P02905; IBD0.
DR InterPro: IPR001882; Biotin_attach.
DR InterPro: IPR000089; Biotin_lipoyl.
DR Pfam; PF00364; biotin_lipoyl.1.
DR PROSITE; PS00188; BIOTIN; 1.
KW Biotin.
SQ SEQUENCE 129 AA; 12687 MW; 205642A39380DCE CRC64;

Query Match 33.3%; Score 205; DB 2; Length 129;
Best Local Similarity 39.4%; Pred. No. 2.8e-09;
Matches 50; Conservative 15; Mismatches 56; Indels 6; Gaps 2;

QY 2 KLTATVNGTAYDVVDVDSK---HENPMGTILFGSGTGAAPAAAGAGAG---KAGSGE 55
DB 3 KFNVTYNGTAIDVEVNEKAAAPAAAPAAAPAAAPAAAPAAAPAAAPVPAAGET 62
QY 56 IPAPLAGVSKILVKEGDTYKAGOTVTLVLEAMKMETEINAPTDGKVEKVLKERDAVGG 115
DB 63 VKAMPKGLISVAVSAGAVAKGETLTLLEAMKMQNEIAAHDAVSEVRISANTVSTG 122
QY 116 QGLIKIRG 122
DB 123 DDWVVLG 129

RESULT 5
054030 PRELIMINARY; PRT; 134 AA.
AC 054030:
DT 01-JUN-1998 (TREMblrel. 06, Created)
DT 01-JUN-1998 (TREMblrel. 06, Last sequence update)
DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)
DE Methylmalonyl-CoA decarboxylase, gamma-subunit (EC 4.1.1.41).
OS Propionigenium modestum.
OC Bacteria; Fusobacteriia; Propionigenium.
OX NCBI_TaxID=2333;
RN [1]
RP SEQUENCE FROM N.A.

RC STRAIN-DSM 2376;
RX MEDLINE=96088990; PubMed=9428714;
RA Bott M., Pfister K., Burda P., Kalbermatter O., Wehlike G.,
RA Dimroth P.;
RT "Methylmalonyl-CoA decarboxylase from Propionigenium modestum: Cloning
RT and sequencing of the structural genes and purification of the enzyme
RT complex."
RL Eur. J. Biochem. 250:590-599(1997).
CC -I- COFACTOR: BIOTIN (BY SIMILARITY).
DR EMBL; AJ002015; CA05139.1; -
DR HSSP; P02905; IBD0.
DR InterPro: IPR001882; Biotin_attach.
DR InterPro: IPR000089; Biotin_lipoyl.
DR Pfam; PF00364; biotin_lipoyl.1.
DR PROSITE; PS00188; BIOTIN; 1.
KW Biotin; lyase.
SQ SEQUENCE 134 AA; 13094 MW; 72CC813187273873 CRC64;

Query Match 32.3%; Score 198.5; DB 2; Length 134;
Best Local Similarity 38.5%; Pred. No. 9.6e-09;
Matches 52; Conservative 12; Mismatches 44; Indels 27; Gaps 3;

QY 4 KTVNGTAYDVVDVDSKSHENPMGTILFGSGTGG-----APAPAG 44
DB 5 KTVNGTETDYDAVE-----EMGGAAYASAPAAAPAAAPAAAPAAAPAPAPAPRTT 57
QY 45 GAGACKAGEGETPAPLAGVSKILVKEGDTYKAGOTVTLVLEAMKMETEINAPTDGKVEK 104
DB 58 AAGAG-AGANVTYAPMPGTTILNVGCHAGDKVSKGDTLVLEAMKMETEINAPHDGVSEV 116
QY 105 LKERDAVGGGGLI 119
DB 117 RVQAGASVNAIGIIV 131

RESULT 6
092AA7 PRELIMINARY; PRT; 145 AA.
AC 092AA7:
DT 01-MAY-1999 (TREMblrel. 10, Created)
DT 01-MAY-1999 (TREMblrel. 10, Last sequence update)
DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)
DE Glutaconyl-CoA decarboxylase gamma subunit (EC 4.1.1.70).
GN GCDC.
OS Acidaminococcus fermentans.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC Clostridiales; Acidaminococcaceae; Acidaminococcus.
OX NCBI_TaxID=905;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 25085;
RX MEDLINE=99157555; PubMed=10027965;
RA Braune A., Bendrat K., Rospert S., Buckel W.;
RT "The sodium ion translocating glutaconyl-CoA decarboxylase from
RT Acidaminococcus fermentans: cloning and function of the genes forming
RT a second operon."
RL Mol. Microbiol. 31:473-487(1999).
CC -I- COFACTOR: BIOTIN (BY SIMILARITY).
DR EMBL; AF030576; AAC69172.1; -
DR HSSP; P02905; IBD0.
DR InterPro: IPR001882; Biotin_attach.
DR InterPro: IPR000089; Biotin_lipoyl.
DR Pfam; PF00364; biotin_lipoyl.1.
DR PRINTS; PR00456; RIBOSOMALP2.
DR PROSITE; PS00188; BIOTIN; 1.
KW Biotin; lyase.
SQ SEQUENCE 145 AA; 13908 MW; 4546006D4F2F4C6B CRC64;

Query Match 32.0%; Score 196.5; DB 2; Length 145;
Best Local Similarity 35.5%; Pred. No. 1.5e-08;
Matches 50; Conservative 14; Mismatches 54; Indels 23; Gaps 2;

```

QY 2 KIKVTYNGTAYDVVD-----VDKSHENPMGTLF-----GGGTGCA 38
DB 3 KEVNVNNGTAYTVEVEGAGATVAPAAAPAAAPAAAPAAAPAAAPAAAPAA 62
QY 39 PAPAAGAGAGAGGEGEIPAPLAGTIVKSGDTYKAGQTVLVEAMKETEINAPTD 98
DB 63 PAPAARPAAAAAGAGSTVSAFPMGKILSVNKRFGDVEAGDVLIIIEAMKONEINAPED 122
QY 99 GKVEKVLVKERDAVGGGGLI 119
DB 123 GTVESEVRNAGDVTATGDVAV 143

RESULT 7
ID 067484 PRELIMINARY: PRT: 655 AA.
AC 067484;
DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
DT 01-AUG-1998 (TREMBLrel. 20, Last annotation update)
DE Private carboxylase C-terminal domain.
GN PYCA OR AQ_1520.
OS Aquifex aeolicus.
OC Bacteria; Aquificae; Aquificales; Aquificaceae; Aquifex.
OX NCBI_TaxID=63363;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VF5;
RX MEDLINE=98196666; PubMed=9537320;
RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
RA Graham D.E., Overbeek R., Sneed M.A., Keller M., Anjay M., Huber R.,
RA Feldman R.A., Short J.M., Olson G.J., Swanson M.V.
RT "The complete genome of the hyperthermophilic bacterium Aquifex
RT aeolicus".
RL Nature 392:353-358(1998).
DR EMBL; AE000744; AAC07445.1; -.
DR HSSP; P02905; 1BDO.
DR InterPro; IPR001882; Biotin_attach.
DR InterPro; IPR00089; Biotin_1lpoyl.
DR InterPro; IPR00379; HMGL_1like.
DR Pfam; PF00364; biotin_1lpoyl; 1.
DR Pfam; PF00682; HMGL_1like; 1.
DR Pfam; PF02436; PYC_OADA; 1.
DR PROSITE; PS00188; BIOTIN; UNKNOWN_1.
KW Pyruvate; Complete proteome.
SQ SEQUENCE 655 AA; 73612 MW; 2839436F6BE05D6 CRC64;

Query Match 31.0%; Score 190.5; DB 16; Length 655;
Best Local Similarity 36.8%; Pred. No. 2.6e-07;
Matches 49; Conservative 19; Mismatches 52; Indels 13; Gaps 3;

QY 2 KIKVTYNGTAYDVVD-----VDKSHENPMGTLF-----GGGTGCA 49
DB 513 KKKVNIIEGVSVAESGKPRKYYRINRLREIQLKPKFAITGGAGQTVGSAEEEGEP 572
QY 50 KAGE-GEIPAPLAGTIVKSGDTYKAGQTVLVEAMKETEINAPTDGVEKVLVE 108
DB 573 KATEPGDVTYPPMPGRKVKILVKEGEPVGGQTVATVEAKMKENEVAPIDGIVTKIFAP 632
QY 109 RDVYGGGGLIKI 121
DB 633 GQVNPDAIMRI 645

RESULT 8
ID 028067 PRELIMINARY: PRT: 140 AA.
AC 028067;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Methylmalonyl-CoA decarboxylase, biotin carboxyl carrier subunit

```

```

DE (MMD).
GN AF2216.
OS Archaeoglobus fulgidus.
OC Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;
OC Archaeoglobaceae; Archaeoglobus.
OX NCBI_TaxID=2234;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
RX MEDLINE=98049343; PubMed=9389475;
RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
RA Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
RA Richardson D.L., Kertavage A.R., Graham D.E., Kyrtides N.C.,
RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA Kirsness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
RA Peterson S., Reich C.I., McNeil L.R., Badger J.H., Glodek A., Zhou L.,
RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,
RA Cotton M.D., Spriggs T., Artlich P., Kaine B.P., Sykes S.M.,
RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA Venter J.C.;
RT "The complete genome sequence of the hyperthermophilic, sulphate-
RT reducing archaeon Archaeoglobus fulgidus.";
RL Nature 390:364-370(1997).
DR EMBL; AE000952; AAB89036.1; -.
DR HSSP; P20708; 1GHJ.
DR TIGR; AF2216; -.
DR InterPro; IPR001882; Biotin_attach.
DR InterPro; IPR00089; Biotin_1lpoyl.
DR Pfam; PF00364; biotin_1lpoyl; 1.
DR PROSITE; PS00188; BIOTIN; UNKNOWN_1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 140 AA; 15686 MW; 30B449C45489C14A CRC64;

Query Match 30.9%; Score 190; DB 17; Length 140;
Best Local Similarity 35.1%; Pred. No. 4.9e-08;
Matches 52; Conservative 19; Mismatches 37; Indels 40; Gaps 4;

QY 1 MKIKVTYNGTAYDVVDVDKSHENPMGTLF-----GGGTGCA 40
DB 1 MKVEYKVGKKRYEVE-----EVSPP---VEYKNGKAAVEVKKKPFKEKADIRE 53
QY 41 -----PANGAGAGAGGEGEIPAPLAGTIVKSGDTYKAGQTVLVEAMKETE 91
DB 54 RFAERREERAREVAKATGKA---ITAPAGVFTILKVKGEKVAAGETVLIIEAMKKN 109
QY 92 EINAPTDGVEKVLVKERDAVGGGGLI 119
DB 110 PIASPDGEIATVVEGDKVASGDVLY 137

RESULT 9
ID 099ZL6 PRELIMINARY: PRT: 116 AA.
AC 099ZL6;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Putative methylmalonyl-CoA decarboxylase, gamma-subunit
DE (EC 4.1.1.41).
GN SPY1176.
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Bacilli; Clostridium group; Lactobacillales;
OC Streptococcaceae; Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SF370 / ATCC 700294 / SEROTYPE M1;
RX MEDLINE=21192684; PubMed=11296296;
RA Ferretti J.J., McShan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K.,
RA Primeaux C., Seate S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P.,
RA Qian Y., Jia H.G., Najjar F.Z., Ren Q., Zhu H., Song L., White J.,
RA Yuan X., Clifton S.W., Roe B.A., McLaughlin R.;

```


GN PH0834.
 OS Pyrococcus horikoshii.
 OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
 OC Pyrococcus.
 NCBI_TaxID=53953;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=OT3;
 RX MEDLINE=98344137; PubMed=9679194;
 RA Kawarabayashi Y., Sawada M., Horikawa H., Halkawa Y., Hino Y.,
 RA Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai Y.,
 RA Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamita M., Ohfuku Y.,
 RA Funahashi T., Tanaka T., Kudoh Y., Yamazaki J., Kushiida N., Oguchi A.,
 RA Aoki K.-I., Yoshizawa T., Nakamura Y., Robb P.T., Horikoshi K.,
 RA Masuchi Y., Shizuya H., Kikuchi H.;
 RT Complete sequence and gene organization of the genome of a hyper-
 thermophilic archaeobacterium, Pyrococcus horikoshii OT3.";
 RL DNA Res. 5:55-76(1998).
 DR EMBL: AP000003; BAA29928.1;
 DR HSSP: P02905; 1BD0.
 DR InterPro: IPR001882; Biotin_attch.
 DR InterPro: IPR000089; Biotin_lipoyl.
 DR InterPro: IPR000891; HMG1-like.
 DR InterPro: IPR003379; PYC_OAD.
 DR Pfam: PF00364; biotin_lipoyl. 1.
 DR Pfam: PF00682; HMG1-like. 1.
 DR Pfam: PF02436; PYC_OAD. 1.
 DR TIGRfams: TIGR01106; oad. 1.
 DR PROSITE: PS00188; BIOTIN; UNKNOWN_1.
 DR Complete proteome.
 SO SEQUENCE 571 AA; 64748 MW; B308B0B1C0E5B103 CRC64;

Query Match 29.9%; Score 184; DB 17; Length 571;
 Best Local Similarity 37.5%; Pred. No. 7.4e-07;
 Matches 45; Conservative 24; Mismatches 43; Indels 8; Gaps 3;

QY 3 LKVTNGTAVDVID-VDKSHENPMGTILFGCGTGAPAPAGAGAGKAGEEIPAPLA 61
 DB 458 IKIYINKEEVEVEGELEFPKPKPOV---QALPSOPPKREYVAPSGSV---VSAFMP 510
 QY 62 GYVSKILVKEGDTVKAQTVLVEAMKETEINAPTDGKYEKVLKERDAVGGGGLIKI 121
 DB 511 GKVLRLVLRVGDVVRVGGGLVLEAMKEMENIPSPRGVVKRLIVKGEAVDVGQPIEL 570

RESULT 13

Q8R7M0 PRELIMINARY; PRT; 122 AA.
 AC Q8R7M0;
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Biotin carboxyl carrier protein.
 GN ACCB2 OR TTE2383.
 OS Thermoanaerobacter tengcongensis.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
 OC Thermoanaerobacteriales; Thermoanaerobacteriaceae; Thermoanaerobacter.
 NCBI_TaxID=119072;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MB4T; JCM11007;
 RX MEDLINE=21992816; PubMed=11997336;
 RA Bao Q., Tian Y., Li W., Xu Z., Xuan Z., Hu S., Dong W., Yang J.,
 RA Chen Y., Xue Y., Xu Y., Lai X., Huang L., Dong X., Ma Y., Ling L.,
 RA Tan H., Chen R., Wang J., Yu J., Yang H.;
 RT A complete sequence of T. tengcongensis genome.";
 RL Genome Res. 12:689-700(2002).
 DR EMBL: AE01180; BAM25522.1;
 DR Complete proteome.
 SO SEQUENCE 122 AA; 13437 MW; 649BFC4629C337C5 CRC64;

Query Match 29.2%; Score 179.5; DB 16; Length 122;
 Best Local Similarity 34.5%; Pred. No. 2.9e-07;

Matches 48; Conservative 12; Mismatches 36; Indels 43; Gaps 3;
 QY 2 LKVTNGTAVDVID-VDKSHENPMGTILFGCGTGAPAPAGAGAGKAGEEIPAPLA 40
 DB 3 KFKVTNGKTYEVEVEEMKAEKEKSEKEIVEYKVPPOEK----- 46
 QY 41 PAAGGAGAGKAGEEIPAPLAGTYSKILVKEGDTVKAQTVLVEAMKETEINAPTDGK 100
 DB 47 -----VSTGR-GSKVVSAPMPGTTIDVRKRGDRKGVLLLEAMKEMENIAPEDGI 100
 QY 101 VEKVLKERDAVGGGGLI 119
 DB 101 VASVNSKASVNTGDVIV 119

RESULT 14

Q8U917 PRELIMINARY; PRT; 576 AA.
 AC Q8U917;
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Biotin carboxylase.
 GN Atu3913 OR AGR_L_1864.
 OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Rhizobiaceae; Rhizobium.
 NCBI_TaxID=176299;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21608550; PubMed=11743193;
 RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitaajima J.P.,
 RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F., Jr., Woo L.,
 RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Boye D., Sr.,
 RA Chapman P., Clendenning J., Deatherage G., Gillist W., Grant C.,
 RA Kutayavin T., Levy R., Li M.-J., McLelland E., Palmeri A.,
 RA Raymond C., Rouse G., Saenphimachak C., Wu Z., Romero P., Gordon D.,
 RA Zhang S., Too H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
 RA Gordon-Ramm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
 RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
 RA Nestor E.W.;
 RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
 C58.";
 RL Science 294:2317-2323(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21608551; PubMed=11743194;
 RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
 RA Qurollo B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,
 RA Houniel K., Gordon J., Vaudin M., Tatchouk O., Epp A., Liu F.,
 RA Williams C., Allinger M., Dougherty D., Scott C., Lappas C., Markelz B.,
 RA Flanagan C., Crowell C., Gurson J., Lomo C., Sear C., Strub G.,
 RA Cleto C., Slater S.;
 RT Genome sequence of the plant pathogen and biotechnology agent
 Agrobacterium tumefaciens C58.";
 RL Science 294:2323-2328(2001).
 DR EMBL: AE009322; ALA44721.1;
 DR EMBL: AE008292; AAK89506.1;
 DR Complete proteome.
 SO SEQUENCE 576 AA; 61722 MW; E0299479A952581F CRC64;

Query Match 28.7%; Score 176.5; DB 16; Length 576;
 Best Local Similarity 43.8%; Pred. No. 3e-06;
 Matches 42; Conservative 17; Mismatches 28; Indels 9; Gaps 3;

QY 26 MGTILFGCGTGAPAPAGAGAGKAGEEIPAPLAGTYSKILVKEGDTVKAQTVLVE 85
 DB 488 LGTV--SGNASAPSAV-----EKKEEMTAAPSGLTQSRVYKDGTVSGDLLAVNE 539
 QY 86 AMKETEINAPTDGKYEKVLKERDAVGGGGLIKI 121
 DB 540 AMKETEINAPTDGKYEKVLKERDAVGGGGLIKI 121

```

RESULT 15
Q9WZH6 PRELIMINARY; PRT: 134 AA.
ID Q9WZH6
AC Q9WZH6
DT 01-NOV-1999 (TREMblrel. 12, Created)
DT 01-NOV-1999 (TREMblrel. 12, Last sequence update)
DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)
DE Propionyl-CoA carboxylase, gamma subunit.
GN TM0717.
OS Thermotoga maritima.
OC Bacteria; Thermotogae; Thermotogaceae; Thermotoga.
OX NCBI_TaxID=2336;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MSB8 / DSM 3109;
RC MEDLINE=99287316; PubMed=10360571;
RA Nelson K.E., Clayton R.A., Gill S.R., Gwin M.L., Dodson R.J.,
RA Halt D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
RA McDonald L., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M.,
RA Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
RA Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
RA Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
RT "Evidence for lateral gene transfer between Archaea and Bacteria from
RT genome sequence of Thermotoga maritima.";
RL Nature 399:323-329(1999).
DR EMBL; AE001743; AAD35799.1; -
DR HSSP; P02905; 1BDO.
DR TIGR; TM0717; -
DR InterPro; IPR000089; Biotin_lipoyl.
DR Pfam; PF00364; biotin_lipoyl; 1.
KW Complete proteome.
SQ SEQUENCE 134 AA; 15131 MW; A3BFBA8574EC1 CRC64;

Query Match 28.6%; Score 176; DB 16; Length 134;
Best Local Similarity 33.8%; Pred. No. 6.1e-07;
Matches 49; Conservative 18; Mismatches 37; Indels 42; Gaps 4;

QY 2 KLTATVNGTAYDQVDV---VDKSH-----ENPMGTIIFGGGTGA 38
DB 4 KRRVYVNGKEYIVEIGNVAKKEPAKEVSQKTVQEIPEKPEPVVL----- 55
QY 39 PAPAAGGAGAGACAGE---IPAPLAGYVKILYKEDTVKAGQVYLVLEAMKMEIENA 95
DB 56 -----EREKSSQOEKLVKAPMAGIVLKYLVKEGQKYNVDKILVEAMKMEIENS 107
QY 96 PTGKVEKYLKERVADYOGGGLIKI 121
DB 108 EFGSTVKELVKEGDNIEIGTILMKI 133

RESULT 16
Q48826 PRELIMINARY; PRT: 596 AA.
ID Q48826
AC Q48826
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE Oxaloacetate decarboxylase alpha-chain.
GN OADA.
OS Legionella pneumophila.
OC Bacteria; Proteobacteria; gamma subdivision; Legionellaceae group;
OC Legionellaceae; Legionella.
OX NCBI_TaxID=446;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CORBY;
RC MEDLINE=97120897; PubMed=8961567;
RA Jahn B., Brand B.C., Lueck P.C., Di Bevardino M., Dimroth P.,
RA Hacker J.;
RT "An oxaloacetate decarboxylase homologue protein influences the
RT intracellular survival of Legionella pneumophila.";
RL FEMS Microbiol. Lett. 145:273-279(1996).

```

```

CC -I- COFACTOR: BIOTIN (BY SIMILARITY).
DR EMBL; X96678; CA67994.1; -
DR HSSP; P11961; 1LAB.
DR InterPro; IPR001882; Biotin_attach.
DR InterPro; IPR000089; Biotin_lipoyl.
DR InterPro; IPR000891; HMGL-1like.
DR InterPro; IPR00379; PYC_OADA.
DR InterPro; IPR000634; S/T_dehydrtse.
DR Pfam; PF00364; biotin_lipoyl; 1.
DR Pfam; PF00682; HMGL-1like; 1.
DR Pfam; PF02436; PYC_OADA; 1.
DR TIGRfams; TIGR01108; oada; 1.
DR PROSITE; PS00188; BIOTIN; 1.
DR PROSITE; PS00165; DEHYDRATASE_SER_THR; 1.
KW Biotin.
SQ SEQUENCE 596 AA; 65667 MW; 55DBEAF96919C86 CRC64;

Query Match 28.6%; Score 176; DB 2; Length 596;
Best Local Similarity 51.4%; Pred. No. 3.4e-06;
Matches 37; Conservative 12; Mismatches 23; Indels 0; Gaps 0;

QY 50 KAGEGIPAPLAGTVSKILYKEDTVKAGQVYLVLEAMKMEIENAPTDGKVEKYLKER 109
DB 522 KIGPGDIYVATPGSTIAIHVSAGDEVKAGQAVLVLEAMKMEIENAPANGVVAETICQKG 581
QY 110 DAVOGGQGLIKI 121
DB 582 DKVTPGQVLIHV 593

RESULT 17
Q9XBJ1 PRELIMINARY; PRT: 984 AA.
ID Q9XBJ1
AC Q9XBJ1
DT 01-NOV-1999 (TREMblrel. 12, Created)
DT 01-NOV-1999 (TREMblrel. 12, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE pyruvate carboxylase (EC 6.4.1.1) (Fragment).
GN PYCA.
OS Bacillus cereus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Bacillaceae; Bacillus.
OX NCBI_TaxID=1396;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 10987;
RC MEDLINE=99231848; PubMed=10217496;
RA Okstad O.A., Hegna I., Lindbaeck T., Rishovd A.L., Kolsto A.B.;
RT "Genome organisation is not conserved between Bacillus cereus and
RT Bacillus subtilis.";
RT Microbiology 145:621-631(1999).
DR EMBL; AJ010111; CAB40604.1; -
DR HSSP; P24182; 1DV2.
DR InterPro; IPR000089; Biotin_lipoyl.
DR InterPro; IPR000901; CPSase.
DR InterPro; IPR000891; HMGL-1like.
DR InterPro; IPR00379; PYC_OADA.
DR Pfam; PF02785; Biotin_carb_C; 1.
DR Pfam; PF00364; biotin_lipoyl; 1.
DR Pfam; PF02786; CPSase_L_D2; 1.
DR Pfam; PF00682; HMGL-1like; 1.
DR Pfam; PF02436; PYC_OADA; 1.
DR PRINTS; PR00098; CPSASE.
DR TIGRfams; TIGR01235; pyruv_carbox; 1.
DR PROSITE; PS00867; CPSASE_2; UNKNONN_1.
KW ligase.
FT NON_TER
SQ SEQUENCE 984 AA; 110135 MW; 7AB53F8D453A147D CRC64;

Query Match 28.6%; Score 176; DB 2; Length 984;
Best Local Similarity 33.8%; Pred. No. 6.1e-06;
Matches 45; Conservative 19; Mismatches 49; Indels 20; Gaps 2;

```

QY 9 GTAYDVVDVDSK-----HENPMGTILFGCGTGAP-----APAGAGAGA 48
 DB 850 GEEDIVEEGKTLMTLVLSIGEPQPGNRLYLEFNGQPREIYVKDESVATVAQRYKG 909
 QY 49 GKAGEGEPAPLACTVSKILVKEGDTVAKAGOTVLEAMKMETEINAPTDGKVEKVLKYE 108
 DB 910 NREPNHISANMPTVTKVYVKEGDEYKKGDSMAITEMAKMETVQAPFNKQKAVYND 969
 QY 109 RDAVQGGQGLIKI 121
 DB 970 GDAIQTGDLIEL 982

RESULT 18
 Q9HPP8
 ID Q9HPP8 PRELIMINARY; PRT; 610 AA.
 AC Q9HPP8;
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Biotin carboxylase.
 GN ACC OR VNG1332G.
 OS Halobacterium sp. (strain NRC-1).
 OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
 OC Halobacteriaceae; Halobacterium.
 NCBI_TaxId=64091;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-20504483; PubMed-11016950;
 RA N.M.V., Kennedy S.P., Mahairas G.G., Bergquist B., Pan M.,
 RA Shukla H.D., Lasky S.R., Baliga N.S., Thorson V., Shroga J.,
 RA Swartzell S., Welt D., Hall J., Dahl T.A., Welti R., Goo Y.A.,
 RA Leilauser B., Keller K., Cruz R., Danson M.J., Hough D.W.,
 RA Leddocks D.G., Jablonaki P.E., Krebs M.P., Angevine C.M., Dale H.,
 RA Isenberger T.A., Peck R.F., Pohlischer M., Spudich J.L., Jung K.-H.,
 RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
 RA Ebhardt H., Lowe T.M., Liang P., Riley M., Hood L., DasSarma S.,
 RA "Genome sequence of Halobacterium species NRC-1."
 RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
 DR EMBL; AEO05066; MAG19819.1; -.
 DR HSSP; P24182; IBNC.
 DR InterPro: IPR001882; Biotin_attach.
 DR InterPro: IPR000089; Biotin_lipoyl.
 DR InterPro: IPR000901; CPSase.
 DR Pfam: PF02785; Biotin_cardc. 1.
 DR Pfam: PF00364; Biotin_lipoyl. 1.
 DR Pfam: PF02786; CPSase_1.D2; 1.
 DR Pfam: PF02786; CPSase_1.D2; 1.
 DR PROSITE: PS00188; BIOTIN; UNKNOWN_1.
 DR PROSITE: PS00867; CPSASE_2; UNKNOWN_1.
 KW Complete proteome.
 SQ SEQUENCE 610 AA; 65805 MW; FC586D84382D6F2 CRC64;

Query Match 28.5%; Score 175.5; DB 17; Length 610;
 Best Local Similarity 38.1%; Pred. No. 3.8e-06;
 Matches 48; Conservative 18; Mismatches 43; Indels 17; Gaps 5;

QY 5 VYNGTAYDVVDVDSKSHENPMGTILFG-----GTGAPAPAGAGAGAGA---GEGE- 55
 DB 492 VEYNGKRFENLE-----ERGAQFAPEADPTGGGPREP-AGGADDETVVEGGET 543
 QY 56 IPAPLAGTYSKILVKEGDTVAKAGOTVLEAMKMETEINAPTDGKVEKVLKRDVAGG 115
 DB 544 VYEMOGTILDAVSEGDADADVLVEAMKRENDVYASHGTYQVAVSEDDSDYMD 603
 QY 116 OGLIKI 121
 DB 604 DVLVVI 609

RESULT 19
 Q8RON4
 ID Q8RON4 PRELIMINARY; PRT; 591 AA.

AC Q8RON4;
 DT 01-JUN-2002 (TREMBLrel. 21, Created)
 DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Acyl-CoA carboxylase B and C subunit.
 GN ACCBC.
 OS Corynebacterium efficiens.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Corynebacterineae; Corynebacteriaceae;
 OC Corynebacterium.
 NCBI_TaxId=152794;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Hirano S., Kimura E., Kawahara Y., Sugimoto S.,
 RT "accBC of Corynebacterium efficiens."
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AB083052; BAB88668.1; -.
 SQ SEQUENCE 591 AA; 63256 MW; 345DEC36C5D8ACA CRC64;

Query Match 28.3%; Score 174; DB 2; Length 591;
 Best Local Similarity 36.3%; Pred. No. 4.9e-06;
 Matches 45; Conservative 20; Mismatches 45; Indels 14; Gaps 3;

QY 2 KLYTVNGTAYDVVDVDSKSHENPMGTILFGCGTGAPAP-----AAGAGAGAGGGEIP 57
 DB 477 KVLVEIDGRVEVALP-----GDLALGGAGAGAKKRRRAAGAGAGVSGD-SVA 526
 QY 58 APPLAGTYSKILVKEGDTVAKAGOTVLEAMKMETEINAPTDGKVEKVLKRDVAGG 117
 DB 527 APAGGTIVKYNVEDGAEVSGDVTYVLEAMKRENDVYASHGTYQVAVSEDDSDYMD 586
 QY 118 LIRI 121
 DB 587 LLEI 590

RESULT 20
 Q99ZL1
 ID Q99ZL1 PRELIMINARY; PRT; 132 AA.
 AC Q99ZL1;
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Putative decarboxylase, gamma chain.
 GN SPY1183.
 OS Streptococcus pyogenes.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
 OC Streptococcaceae; Streptococcus.
 NCBI_TaxId=1314;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN-SF370 / ATCC 700294 / SEROTYPE M1;
 RC MEDLINE-21192684; PubMed-11296296;
 RA Ferretti J.J., McShan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K.,
 RA Primeaux C., Sezate S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P.,
 RA Qian Y., Jia H.G., Najjar F.Z., Ren Q., Zhu H., Song L., White J.,
 RA Yuan X., Clifton S.W., Roe B.A., McLaughlin R.,
 RT "Complete genome sequence of an M1 strain of Streptococcus pyogenes."
 RL Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).
 DR EMBL; AEO06559; AAK34049.1; -.
 DR HSSP; P02905; 3BD0.
 DR InterPro: IPR001882; Biotin_attach.
 DR InterPro: IPR000089; Biotin_lipoyl.
 DR Pfam: PF00364; Biotin_lipoyl. 1.
 DR PROSITE: PS00188; BIOTIN; 1.
 KW Biotin; Complete proteome.
 SQ SEQUENCE 132 AA; 13529 MW; 2FA2B1FB8119429E CRC64;

Query Match 28.2%; Score 173.5; DB 16; Length 132;
 Best Local Similarity 32.4%; Pred. No. 9.5e-07;
 Matches 47; Conservative 15; Mismatches 42; Indels 41; Gaps 2;

RL J Bacteriol. 184:2005-2018(2002).
DR EMBL: AE010533; AL944406.1; -
KW Lyase; Complete proteome.
SQ SEQUENCE 134 AA; 14063 MW; 177943f669196AC6 CRC64;
Query Match 28.0%; Score 172; DB 16; Length 134;
Best Local Similarity 34.8%; Pred. No. 1.3e-06;
Matches 46; Conservative 20; Mismatches 52; Indels 14; Gaps 2;
QY 1 MKLAVTVNGTAYDVVD-----VDKSHENPMGTILFGGTGAPAPAGAG 47
DB 1 MKYVTVNGKFEVEVEVGAGKSLRQPYERETVYKSPVETKVAAPVAAAPAT 60
QY 48 AGKAGEEIPRLAGTYSKILYKGDYKAGQYLVLEAMKMEIINAPDGYEKVLYK 107
DB 61 ATTGGT-TITSPMGSTILDYKVNVDKVFQGTILALEAMKMEIINPATADGEVAFIRVK 119
QY 108 ERDAVGGGGLIKI 119
DB 120 KGDVETDSVLT 131
RESULT 24
P71122 PRELIMINARY; PRT; 591 AA.
AC P71122;
DT 01-FEB-1997 (TEMBLrel. 02, Created)
DT 01-FEB-1997 (TEMBLrel. 02, Last sequence update)
DT 01-MAR-2002 (TEMBLrel. 20, Last annotation update)
DE Acyl coenzyme A carboxylase.
GN ACCBC.
OS Corynebacterium glutamicum (Brevibacterium flavum).
OC Bacteria; Firmicutes; Actinobacteriia; Actinobacteriales;
OC Actinomycetales; Corynebacterineae; Corynebacteriaceae;
OC Corynebacterium.
OX NCBI_TaxID=1718;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 13032;
RX MDL:LINE-96337861; PubMed-8772169;
RA Jager W., Peters-Wendisch P.G., Kalinowski J., Puhler A.;
RT "A Corynebacterium glutamicum gene encoding a two-domain protein
RT similar to biotin carboxylases and biotin-carboxyl-carrier proteins.";
RL Arch. Microbiol. 166:76-82(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 13032;
RA Jeeger W.;
RL Submitted (MUG-1995) to the EMBL/GenBank/DBJ databases.
CC -1- COFACTOR: BIOTIN (BY SIMILARITY).
DR EMBL: U03023; AAB40890.1; -
DR HSSP; P24182; IBCN.
DR InterPro: IPR001882; Biotin_attach.
DR InterPro: IPR000089; Biotin_lipoyl.
DR InterPro: IPR000901; CPSase.
DR Pfam: PF02785; Biotin_card_C. 1.
DR Pfam: PF00364; Biotin_lipoyl. 1.
DR Pfam: PF00289; CPSase_L_D2; 1.
DR Pfam: PF02786; CPSase_L_D2; 1.
DR PROSITE: PS00188; Biotin. 1.
DR PROSITE: PS00867; CPSASE_2; UNKNOWN_1.
KW Biotin.
SQ SEQUENCE 591 AA; 63419 MW; 0C9ED9DE184F8E84 CRC64;
Query Match 27.7%; Score 170.5; DB 2; Length 591;
Best Local Similarity 38.7%; Pred. No. 9.3e-06;
Matches 48; Conservative 16; Mismatches 45; Indels 15; Gaps 3;
QY 2 KLAFTVNGTAYDVVDVVDKSHENPMGTILFGGTGAPAPAGAG 57
DB 478 KVAVEINGRVEVALPD-----LALGTAAGPKKRAKRAAGACAKAGVSGDA-VA 526
QY 58 APLAGTYSKILYKGDYKAGQYLVLEAMKMEIINAPDGYEKVLYKERDAVGGG 117

DB 527 APMQGTIVKYNVEGAEVNEEDTVVLEAMKMEIINAPDGYEKVLYK 586
QY 118 LIKI 121
DB 587 LIKI 590
RESULT 25
Q97ER7 PRELIMINARY; PRT; 1144 AA.
AC Q97ER7;
DT 01-OCT-2001 (TEMBLrel. 18, Created)
DT 01-OCT-2001 (TEMBLrel. 18, Last sequence update)
DT 01-JUN-2002 (TEMBLrel. 21, Last annotation update)
DE Pyruvate carboxylase, PYKA.
GN CAC2660.
OS Clostridium acetobutylicum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC Clostridiales; Clostridiaceae; Clostridium.
OX NCBI_TaxID=1488;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 824 / DSM 792 / VKM B-1787;
RX MDL:LINE-21359325; PubMed-11466286;
RA Noelling J., Breton G., Omelchenko M.V., Makarova K.S., Zeng Q.,
RA Gibson R., Lee H.M., Dubois J., Qiu D., Hitt J., Wolf Y.I.,
RA Tatusov R.L., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.J.,
RA Bennett G.N., Koehn E.V., Smith D.R.;
RT "genome sequence and comparative analysis of the solvent-producing
RT bacterium Clostridium acetobutylicum";
RL J. Bacteriol. 183:4823-4838(2001).
DR EMBL: AE007763; AAK80607.1; -
DR InterPro: IPR001882; Biotin_attach.
DR InterPro: IPR000089; Biotin_lipoyl.
DR InterPro: IPR000901; CPSase.
DR InterPro: IPR000891; HMG-Like.
DR InterPro: IPR003141; PHP_N.
DR InterPro: IPR003379; PYC_OADA.
DR Pfam: PF02785; Biotin_card_C. 1.
DR Pfam: PF00364; Biotin_lipoyl. 1.
DR Pfam: PF00289; CPSase_L_D2; 1.
DR Pfam: PF02786; CPSase_L_D2; 1.
DR Pfam: PF00682; HMG-Like. 1.
DR Pfam: PF02436; PYC_OADA. 1.
DR PRINTS; PR00098; CPSASE.
DR SMART; SM00481; POLIITAC. 1.
DR TIGR: TIGR01235; Pyruv_carbox. 1.
DR PROSITE: PS00188; BIOTIN; UNKNOWN_1.
DR PROSITE: PS00866; CPSASE_1; 1.
DR PROSITE: PS00867; CPSASE_2; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 1144 AA; 127709 MW; 519FA29A8008F326 CRC64;
Query Match 27.6%; Score 170; DB 16; Length 1144;
Best Local Similarity 47.3%; Pred. No. 2.2e-05;
Matches 35; Conservative 13; Mismatches 26; Indels 0; Gaps 0;
QY 48 AGKAGEEIPRLAGTYSKILYKGDYKAGQYLVLEAMKMEIINAPDGYEKVLYK 107
DB 1070 ADSNNKKEIASIPGNVYKVFAPGDKGDSLVYIAKMEIINVSSEDEGVGIFVK 1129
QY 108 ERDAVGGGGLIKI 121
DB 1130 EGDVVOGGLVLT 1143
RESULT 26
Q9HTD1 PRELIMINARY; PRT; 607 AA.
AC Q9HTD1;
DT 01-MAR-2001 (TEMBLrel. 16, Created)
DT 01-MAR-2001 (TEMBLrel. 16, Last sequence update)

DT 01-JUN-2002 (TREMBLrel. 21, last annotation update)
DE Probable transcarboxylase subunit.
GN PA5435.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
CC Pseudomonas.
NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PA01;
RX MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Pham X.-O.T., Erwin A.L., Mizoguchi S.D., Warrenner P., Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M., Gardner R.L., Goltz L., Tolentino E., Westbrock-Wadman S., Yuan Y., Brody L.L., Coulter S.N., Folger K.R., Kas A., Lardig K., Lim R.M., Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T., Ralzer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
RA "Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen."
RT Nature 406:959-964(2000).
RL EMBL; AE004956; AAC08820.1; -
DR HSSP; P02905; 3BDO.
DR InterPro; IPR000089; Biotin_lipoyl.
DR InterPro; IPR000891; HMG1-like.
DR InterPro; IPR003379; PYC_OXA.
DR InterPro; IPR000634; S/T_dehydratase.
DR Pfam; PF00364; Biotin_lipoyl; 1.
DR Pfam; PF00682; HMG1-like; 1.
DR Pfam; PF02436; PYC_OXA; 1.
DR TIGRfams; TIGR01108; oxa; 1.
DR PROSITE; PS00165; DEHYDRATASE_SER_THR; UNKNOWN.1.
RW Complete proteome.
SQ SEQUENCE 607 AA; 66095 MW; 232AB0E9B935E010 CRC64;

Query Match 27.4%; Score 168.5; DB 16; Length 607;
Best Local Similarity 32.4%; Pred. No. 1.4e-05;
Matches 44; Conservative 20; Mismatches 43; Indels 29; Gaps 3;

QY 5 VTNGTAVDVVDV-----DVDKSHE---NPMGTILFGGGTGAPAPAAAG 45
DB 480 IDVHGTYAVDITGVKSDNKRHFYLSIDGMPVEVFLENYVAGSAGKHA----- 534
QY 46 AGAGNAGEEIPAPLAGYVKILVEGDTVAKGQTVLYLEAKMETEINAPDQKVE 105
DB 535 -----SEPHVSTPMGNVVDLVKESGVKAGCAVLLTEAKMETEVQAGIAGYKAIH 589
QY 106 VKERDVGOGGGLIKI 121
DB 590 VAKGDVNPGEILIELI 605

RESULT 27
067544 PRELIMINARY; PRT; 620 AA.
AC 067544;
DT 01-AUG-1998 (TREMBLrel. 07, Created)
DT 01-AUG-1998 (TREMBLrel. 07, last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, last annotation update)
DE Oxaloacetate decarboxylase alpha chain.
GN OXA OR AO.1614.
OS Aquifex aeolicus.
OC Bacteria; Aquificae; Aquificales; Aquificaceae; Aquifex.
NCBI_TaxID=63363;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VF5;
RX MEDLINE=98196666; PubMed=9537320;
RA Decourt G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L., Grahman D.E., Overbeek R., Sneed M.A., Keller M., AuJay M., Huber R., Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
RA "The complete genome of the hyperthermophilic bacterium Aquifex aeolicus".
RT Nature 392:353-358(1998).
RL Nature 392:353-358(1998).

CC -1- FUNCTION: TUBULIN IS THE MAJOR CONSTITUENT OF MICROTUBULES. IT
CC BINDS TWO MOLES OF GTP, ONE AT AN EXCHANGEABLE SITE ON THE BETA
CC CHAIN AND ONE AT A NONEXCHANGEABLE SITE ON THE ALPHA-CHAIN (BY
CC SIMILARITY).
CC -1- COPFACTOR: BIOTIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE TUBULIN FAMILY.
DR EMBL; AE000747; AAC07497.1; -
DR HSSP; P02905; 1BDO.
DR InterPro; IPR002453; Beta_tubulin.
DR InterPro; IPR001882; Biotin_attach.
DR InterPro; IPR000089; Biotin_lipoyl.
DR InterPro; IPR000891; HMG1-like.
DR InterPro; IPR003379; PYC_OXA.
DR Pfam; PF00364; Biotin_lipoyl; 1.
DR Pfam; PF00682; HMG1-like; 1.
DR Pfam; PF02436; PYC_OXA; 1.
DR TIGRfams; TIGR01108; oxa; 1.
DR PROSITE; PS00188; BIOTIN; 1.
DR PROSITE; PS00228; TUBULIN_B_AUTOREG; 1.
RW Biotin; GTP-binding; Microtubules; Complete proteome.
SQ SEQUENCE 620 AA; 70416 MW; 4F306D48794AE859 CRC64;

Query Match 27.4%; Score 168.5; DB 16; Length 620;
Best Local Similarity 33.8%; Pred. No. 1.4e-05;
Matches 47; Conservative 17; Mismatches 52; Indels 23; Gaps 3;

QY 5 VTNGTAVDVVDVVDKSHENPMGTILFGGGTG-----APAPAGCA 46
DB 475 VYVHGQYHVOI-AKGEEDPRFRFFIRLDSQLOEVLQPIREVSVEAFATEBGT 533
QY 47 GAGK-----AGEEIPAPLAGYVKILVEGDTVAKGQTVLYLEAKMETEINAPDQKVE 102
DB 534 VYSKRPKPGKIDVSPITGKVNIVNVGDEVKEGDVLYLEAKMETEINHSVDIIVE 593
QY 103 KVLKERNVAVOGGGLIKI 121
DB 594 ELFVRGTVNPDEVILRI 612

RESULT 28
08RAJ2 PRELIMINARY; PRT; 135 AA.
ID 08RAJ2;
AC 08RAJ2;
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, last annotation update)
DE Biotin carboxyl carrier protein.
GN ACCB OR TBE122.
OS Thermotogaobacter tengcongensis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC Thermotogaobacteriales; Thermotogaobacteriaceae; Thermotogaobacter.
NCBI_TaxID=119072;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MB4T / JCM11007;
RX MEDLINE=21992816; PubMed=11997336;
RA Bao Q., Tian Y., Li W., Xu Z., Xuan Z., Hu S., Dong W., Yang J., Chen Y., Xue Y., Xu Y., Lai X., Huang L., Dong X., Ma Y., Ling L., Tan H., Chen R., Wang J., Yu J., Yang H.;
RA "A complete sequence of T. tengcongensis genome."
RL Genome Res. 12:689-700(2002).
DR EMBL; AE013084; AAM24452.1; -
RW Complete proteome.
SQ SEQUENCE 135 AA; 15195 MW; 9C2BF717C008800A CRC64;

Query Match 27.3%; Score 168; DB 16; Length 135;
Best Local Similarity 30.6%; Pred. No. 2.7e-06;
Matches 45; Conservative 18; Mismatches 38; Indels 46; Gaps 3;

QY 2 KKKVYVNGTAVDVVDV-----VDKSHENPMGTILFGGGTGAPAPAGGAGAKA-- 51
DB 3 KKKVYVNGTAVVLEIEIIOENKTTKEVSEKSPQVLL-----SEKVEQ 45

GenCore version 5.1.4-p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 1, 2003, 07:49:49 ; Search time 10.8021 Seconds

(without alignments)
468.439 Million cell updates/sec

Title: US-09-987-485-1

Perfect score: 615
Sequence: 1 MKLKVTVNGTAYDVVDVVK.....KVLKERDAYGGGGLIKIG 122

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	ID	Description
1	604.5	98.3	123 1	BCCP_PROFR
2	200.5	32.6	595 1	DCOA_KLEPN
3	175.5	28.5	590 1	DCOA_SALRY
4	173	28.1	567 1	PYCB_MERJA
5	168	27.3	130 1	BCCP_STRMU
6	163	26.5	598 1	BCCA_MYCLE
7	159	25.9	717 1	MCCA_MOUSE
8	153	24.9	725 1	MCCA_HUMAN
9	143	23.3	70 1	BCCP_LYCES
10	143	23.3	734 1	MCCA_ARATH
11	141	22.9	568 1	PYCB_MERTH
12	138.5	22.5	157 1	BCCP_PORPU
13	138	22.4	1180 1	PYCB_YEAST
14	133.5	21.7	163 1	BCCP_CHLMU
15	133	21.6	1178 1	PYCB_YEAST
16	132	21.5	1189 1	PYCB_PICPA
17	131	21.3	164 1	BCCP_CHLFR
18	130	21.1	629 1	ODP2_ECOLI
19	130	21.1	731 1	MCCA_SOYBN
20	129.5	21.1	181 1	BCCP_ANASP
21	129	21.0	547 1	ODP2_PSENE
22	129	21.0	1178 1	PYCB_MOUSE
23	128	20.8	544 1	ODP2_ACHLA
24	125.5	20.4	654 1	BCCA_MYCTU
25	123.5	20.1	637 1	ODP2_AZOVI
26	122	19.8	553 1	ODP2_ALCEU
27	121	19.7	1178 1	PYCB_HUMAN
28	119	19.3	704 1	PCCA_RAT
29	118.5	19.3	70 1	BCCP_MYCSM
30	118	19.2	156 1	BCCP_PSEAE
31	117	19.0	70 1	BCCP_PSEAE
32	115	18.7	262 1	BCCP_SOYBN
33	115	18.7	262 1	BCCP_SOYBN

34	114	18.5	438 1	ODP2_MYCCA
35	112	18.2	155 1	BCCP_HAEIN
36	110	17.9	703 1	PCCA_HUMAN
37	108.5	17.6	167 1	BCCP_CHLPN
38	108	17.6	553 1	ODP2_MYCTU
39	107	17.4	70 1	BTR7_MYCLE
40	105.5	17.2	567 1	ODP2_HAEIN
41	104	16.9	1835 1	DUR1_YEAST
42	103	16.7	462 1	ODPB_ZYMO
43	102.5	16.7	156 1	BCCP_ECOLI
44	97	15.8	384 1	ODP2_MYCGE
45	95.5	15.5	463 1	ODP2_YEAST
46	94.5	15.4	152 1	BCCP_YEAST
47	94	15.3	447 1	ODP2_RHIME
48	93.5	15.2	2273 1	HEAL_YEAST
49	93	15.1	402 1	ODP2_MYCPN
50	93	15.1	440 1	ODP2_ZYMO
51	92	15.0	427 1	ODP2_BACST
52	91	14.8	398 1	ACOC_BACSU
53	91	14.8	401 1	ODP2_RICPR
54	91	14.8	417 1	ODP2_BACSU
55	90.5	14.7	159 1	BCCP_ARATH
56	90	14.6	280 1	BCCP_BACSU
57	89.5	14.6	323 1	JUND_CHICK
58	89.5	14.6	1122 1	ADP1_MYCGA
59	88	14.3	430 1	ODP2_STAPH
60	87.5	14.2	1147 1	MYSB_ACACA
61	86.5	14.1	439 1	XP2_XENLA
62	84	13.7	460 1	ODPB_RHIME
63	84	13.7	1733 1	VNUA_PRYKA
64	83.5	13.6	2131 1	CINA_DROME
65	83	13.5	641 1	EBN1_EBV
66	82.5	13.4	355 1	YHII_ECOLI
67	82.5	13.4	778 1	Y034_MYCTU
68	82	13.3	501 1	ODPB_HUMAN
69	82	13.3	1038 1	RSGL_RAT
70	81.5	13.3	420 1	ODP2_BUCAT
71	81.5	13.3	488 1	FLAB_PSENE
72	79.5	12.9	232 1	RL1_PROVU
73	79.5	12.9	398 1	ODP2_AZOVI
74	79	12.8	409 1	ODP2_HUMAN
75	79	12.8	2280 1	COAC_SCHPO
76	78.5	12.8	380 1	VASP_HUMAN
77	78	12.7	231 1	RL1_BACSU
78	78	12.7	453 1	ODP2_FUGRI
79	78	12.7	455 1	ODP2_HUMAN
80	78	12.7	497 1	FXD2_HUMAN
81	78	12.7	599 1	VG9_BPZA
82	78	12.7	905 1	HXAI_HAEIN
83	78	12.7	1861 1	APU_THERU
84	77.5	12.6	233 1	RL1_ECOLI
85	77.5	12.6	233 1	RL1_SERMA
86	77.5	12.6	409 1	ODP2_HAEIN
87	77.5	12.6	424 1	ODP2_BACSU
88	77.5	12.6	441 1	ODP2_SCHPO
89	77.5	12.6	452 1	ODP2_SCHPO
90	77.5	12.6	555 1	ODP2_RAT
91	77	12.5	412 1	MYRC_NEIGO
92	77	12.5	575 1	GGT_PSESP
93	77	12.5	2483 1	COA2_HUMAN
94	76.5	12.4	541 1	YOE8_ECOLI
95	76.5	12.4	819 1	RNFC_HAEIN
96	76.5	12.4	2038 1	FSH_DROME
97	76	12.4	384 1	VASP_CANPA
98	76	12.4	843 1	ARM_DROME
99	76	12.4	3591 1	FHAB_BORPE
100	75.5	12.3	954 1	FLEY_CAUCR

ALIGNMENTS

Q49110	mycoplasma
P43874	haemophilus
P05165	homo sapien
Q92901	chlamydia p
O10381	mycobacteri
Q9cch9	mycobacteri
P5118	haemophilus
P32528	saccharomyc
O66113	zymomonas m
P47514	mycoplasma
P1962	saccharomyc
O19918	cyamidium c
Q39873	rhizobium m
P32874	saccharomyc
P75392	mycoplasma
O66119	zymomonas m
P11961	bacillus st
O31550	bacillus su
Q32d43	rickettsia
P16263	bacillus su
P49766	bacillus su
Q42533	arabidopsis
P27921	gallus galli
Q49379	mycoplasma
Q39821	staphylococ
P19706	acanthamoeb
P17437	xenopus lae
O97944	rhizobium m
P33465	pseudorabie
P35500	drosophila
P03211	epstein-bar
P76526	eschcherichia
P71933	mycobacteri
O00330	homo sapien
P50904	rattus norv
P57389	buchnera ap
P72151	pseudomonas
P10054	proteus vul
P20708	azotobacter
Q90512	figu rubrip
P78820	schizosacch
P50552	homo sapien
Q06797	bacillus su
P36957	homo sapien
Q9n0f1	sus scrofa
O60548	homo sapien
P07534	bacteriophag
P46602	haemophilus
P38536	t amylopull
P02384	eschcherichia
P09764	serattia ma
P45302	haemophilus
P37942	bacillus su
P21883	bacillus su
O54681	schizosacch
P08461	rattus norv
P43505	neisseria g
P36267	pseudomonas
O00767	homo sapien
Q46808	eschcherichia
P13797	haemophilus
P313709	drosophila
P50551	canis famli
P18824	drosophila
P12253	bordeletia
P15345	caulobacter

RESULT 1

BCCP_PROFR
ID BCCP_PROFR STANDARD; PRT; 123 AA.
AC P02904;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Biotin-carboxyl carrier protein of methylmalonyl-CoA carboxyl-
transferase (transcarboxylase, 1.3S subunit).
OS Propionibacterium freudenreichii shermanii.
OC Bacteria: Actinobacteria; Actinobacteriales; Actinobacteriaceae;
OC Actinomycetales; Propionibacteriales; Propionibacteriaceae;
OC Propionibacterium.
OX NCBI_TaxID=1752;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85298212; PubMed=3898065;
RA Murtif V.L., Bahler C.R., Samols D.;
RT "Cloning and expression of the 1.3S biotin-containing subunit of
transcarboxylase.";
RT Proc-Matl Acad Sci U.S.A. 82:5617-5621(1985).
RN [2]
RP SEQUENCE.
RX MEDLINE=80049796; PubMed=40985;
RA Maloy W.L., Bowien B.U., Zwolinski G.K., Kumar G.K., Wood H.G.,
RT Ericsson L.H., Walsh K.A.;
RT "Amino acid sequence of the biotinyl subunit from transcarboxylase.";
RT J. Biol. Chem. 254:11615-11622(1979).
RN [3]
RP MUTAGENESIS OF ALA-87; MET-88; LYS-89 AND MET-90.
RX MEDLINE=92406744; PubMed=1526981;
RA Shenoy B.C., Xie Y., Park V.L., Kumar G.K., Beegen H., Wood H.G.,
RT Samols D.;
RT "The importance of methionine residues for the catalysis of the
biotin enzyme, transcarboxylase. Analysis by site-directed
mutagenesis.";
RT J. Biol. Chem. 267:18407-18412(1992).
RN [4]
RP STRUCTURE BY NMR.
RX MEDLINE=98066416; PubMed=9398186;
RA Reddy D.V., Shenoy B.C., Carey P.R., Soennichsen F.D.;
RT "Absence of observable biotin-protein interactions in the 1.3S
subunit of transcarboxylase: an NMR study.";
RT Biochemistry 36:14676-14682(1997).
RN [5]
RP FUNCTION: THE BIOTINYL 1.3S SUBUNIT SERVES AS A CARBOXYL CARRIER
BETWEEN THE SUBSTRATE BINDING SITES ON THE 12S AND 5S SUBUNITS.
CC -1- SUBUNIT: TRANSCARBOXYLASE IS COMPOSED OF THREE SUBUNITS: 1.3S,
5S, AND 12S. THE CORE OF THE ENZYME IS COMPOSED OF SIX 12S
SUBUNITS. ON EACH SIDE OF THE CORE THERE ARE THREE PAIRS OF 5S
SUBUNITS. EACH 5S DIMER IS ATTACHED TO THE CORE BY TWO 1.3S
SUBUNITS. THUS THE TOTAL NUMBER OF CHAINS IS 30 (6 + 12 + 12).
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: M11738; AAA25674.1; -
DR PIR: A03401; BKIP.
DR HSSP: P02905; 1A6X.
DR InterPro: IPR001882; Biotin_attach.
DR InterPro: IPR000089; Biotin_lipoyl.
DR Pfam: PF00364; biotin_lipoyl; 1.
DR PROSITE: PS00188; BIOTIN; 1.
KW Biotin
FT BINDING 89 89 BIOTIN.
SQ SEQUENCE 123 AA; 12367 MW; D0980CC2065EA9A89 CRC64;
Query Match 98.3%; Score 604.5; DB 1; Length 123;
Best Local Similarity 99.2%; Pred. No. 5, 7e-40;
Matches, 122; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

OY 1 MKLKVYNGTAYDVVDVDSKSHENPMGTLIFGGGTGAPAP-AAAGAGAGKAGEGEIPAP 59
DB 1 MKLKVYNGTAYDVVDVDSKSHENPMGTLIFGGGTGAPAPRAAGAGAGKAGEGEIPAP 60
OY 60 LAQVSKILVKEEDDYKAGQTVLYLEAMKMETEINAPTDGKVKVLYKREDAVQGGGL 119
DB 61 LAQVSKILVKEEDDYKAGQTVLYLEAMKMETEINAPTDGKVKVLYKREDAVQGGGL 120
OY 120 KIG 122
DB 121 KIG 123

RESULT 2
ID COCA_KLEPN STANDARD; PRT; 595 AA.
AC P13187;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 01-JUL-1993 (Rel. 26, Last annotation update)
DE Oxaloacetate decarboxylase alpha chain (EC 4.1.1.3).
GN OADA.
OS Klebsiella pneumoniae.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Klebsiella.
OX NCBI_TaxID=573;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88257085; PubMed=2454915;
RA Schwarz E., Oesterheld D., Reinke H., Beyreuther K., Dimroth P.;
RT "The sodium ion translocating oxaloacetate decarboxylase of Klebsiella
pneumoniae. Sequence of the biotin-containing alpha-subunit and
RT relationship to other biotin-containing enzymes.";
RT J. Biol. Chem. 263:9640-9645(1988).
RN [2]
RP FUNCTION: LYASE AND SODIUM TRANSPORTER.
CC -1- CATALYTIC ACTIVITY: Oxaloacetate - pyruvate + CO(2).
CC -1- COFACTOR: BIOTIN AND REQUIRES A SODIUM ION.
CC -1- SUBUNIT: COMPOSED OF THREE CHAINS (ALPHA, BETA, AND GAMMA).
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: J03885; AAA25120.1; -
DR PIR: A28088; A28088.
DR HSSP: P02905; 1A6X.
DR InterPro: IPR001882; Biotin_attach.
DR InterPro: IPR000089; Biotin_lipoyl.
DR InterPro: IPR000891; HMGL-like.
DR InterPro: IPR003379; PTC-OADA.
DR Pfam: PF00364; biotin_lipoyl; 1.
DR Pfam: PF00682; HMGL-like; 1.
DR Pfam: PF02436; PTC-OADA; 1.
DR TIGRFAMs: TIGR01108; oada; 1.
DR PROSITE: PS00188; BIOTIN; 1.
KW Decarboxylase; Lyase; Sodium transport; Biotin.
FT INIT_MET 561 561
FT BINDING 0 0
SQ SEQUENCE 595 AA; 63402 MW; AA744A9556E9488C CRC64;
Query Match 32.6%; Score 200.5; DB 1; Length 595;
Best Local Similarity 52.9%; Pred. No. 1, 9e-08;
Matches 45; Conservative 10; Mismatches 29; Indels 1; Gaps 1;
OY 38 AAPAPAGAGAGKAGEG-EIPAPLAGTVSKILVKEEDDYKAGQTVLYLEAMKMETEINAP 96
DB 510 AAPAPAPAPAPAPAGAGTVTAPPLACTIKVYLASEQIVYAAAGVLLILEAMKMETEIRRA 569

Qy 97 TDGKVKYKRDVAVGCGGLIKI 121
 Db 570 QAGTVRGIAVAGDVAAGDTLMTL 594

RESULT 3

DCOA_SALTY STANDARD; PRT; 590 AA.
 ID DCOA_SALTY STANDARD; PRT; 590 AA.
 AC 003030;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Oxaloacetate decarboxylase alpha chain (EC 4.1.1.3).
 GN (OADA OR OADA OR STM0055) AND (OADA2 OR STM3352).
 OS Salmonella typhimurium.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 CC Salmonella.
 CX NCBI_TaxID-602;
 RX MEDLINE-3054591; PubMed-1331067;
 RA Woshke G., Wifling K., Dimroth P.;
 RT "Sequence of the sodium ion pump oxaloacetate decarboxylase from
 Salmonella typhimurium";
 RL J. Biol. Chem. 267:22798-22803(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-LT2;
 RX MEDLINE-21534948; PubMed-11677609;
 RA McCelland M., Sanderson K.E., Splith J., Clifton S.W., Latreille P.,
 RA Courtney L., Portwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
 RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
 RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
 RA Waterston R., Wilson R.K.;
 RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
 LT2";
 RL Nature 413:852-856(2001).
 CC -1- FUNCTION: LYSASE AND SODIUM TRANSPORTER.
 CC -1- CATALYTIC ACTIVITY: Oxaloacetate + pyruvate + CO(2).
 CC -1- COFACTOR: BIOTIN AND REQUIRES A SODIUM ION.
 CC -1- SUBUNIT: COMPOSED OF THREE CHAINS (ALPHA, BETA, AND GAMMA).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 or send an email to license@sib-sib.ch).
 DR EMBL; M96434; AAA02973.1;
 DR EMBL; AE008696; AAL19019.1;
 DR EMBL; AE008854; AAL22221.1;
 DR PIR; B44465; B44465.
 DR HSP; P02905; IABX.
 DR StyGene; SG10259; cadal.
 DR InterPro; IPR001882; Biotin_attach.
 DR InterPro; IPR000089; Biotin_lipoyl.
 DR InterPro; IPR000891; HMG-L-like.
 DR Pfam; PF00364; Biotin_lipoyl; 1.
 DR Pfam; PF00682; HMG-L-like; 1.
 DR Pfam; PF02436; HMG-L-like; 1.
 DR TIGRfam; TIGR01108; cadal; 1.
 DR PROSITE; PS00186; BIOTIN; 1.
 DR Decarboxylase; Lyase; Sodium transport; Biotin; Complete proteome.
 FT INT MET 0
 FT BINDING 556 556 BY SIMILARITY.
 FT SEQUENCE 590 AA; 63075 MW; 4EA421F9324AFD7B C6C64;
 Query Match 28.5%; Score 175.5; DB 1; Length 590;

Best Local Similarity 37.5%; Pred. No. 1.5e-06;
 Matches 48; Conservative 10; Mismatches 41; Indels 29; Gaps 3;

Oy 5 TVVNGTAVDVIDVDKSHENPMGCTILFGG-----TGAPAPAGAGAGKAGE 53
 Db 480 VEVEKAFVVKVS-----DGDISQITAAVPAASAPVQAAAPAGAGT--- 522
 Oy 54 GEIPAPLAGTVSKILVREGDTYKAGQTVLYEAMKETEINPTDGKVKYKRDVAVQ 113
 Db 523 -PVTAFLAGNIWKVATISGQTVABGDVLLIEAMKETEINPTDGKVKYKRDVAVQ 581
 Oy 114 GGCGGLIKI 121
 Db 582 VGDITLMTL 589

RESULT 4

PCB_METUA STANDARD; PRT; 567 AA.
 ID PCB_METUA STANDARD; PRT; 567 AA.
 AC 058628;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Pyruvate carboxylase subunit B (EC 6.4.1.1) (Pyruvic carboxylase B).
 GN PCB OR M01231.
 OS Methanococcus jannaschii.
 OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
 CC Archaea; Euryarchaeota; Methanococci; Methanococcales;
 CX NCBI_TaxID-2190;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-JAL-1 / DSM 2661 / ATCC 43067;
 RX MEDLINE-96337999; PubMed-8688087;
 RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
 RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
 RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
 RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
 RA Scott J.L., Geoghegan N.S.M., Feldman J.F., Furumann J.L., Nguyen D.,
 RA Uitterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
 RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
 RA Klenk H.-P., Fraser C.M., Smith H.O., Weese C.R., Venter J.C.;
 RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
 jannaschii";
 RL Science 273:1058-1073(1996).
 RN [2]
 RP SEQUENCE OF 190-125; 260-270; 277-289; 277-289; 309-325; 328-358;
 RP 370-380; 386-409; 422-438; 491-506 AND 491-506, AND FUNCTION.
 RX MEDLINE-21034791; PubMed-11195096;
 RA Mukhopadhyay B., Patel V.J., Wolfe R.S.;
 RT "A stable archaeal pyruvate carboxylase from the hyperthermophile
 Methanococcus jannaschii";
 RL Arch. Microbiol. 174:406-414(2000).
 CC -1- FUNCTION: PYRUVATE CARBOXYLASE CATALYZES A 2-STEP REACTION,
 INVOLVING THE ATP-DEPENDENT CARBOXYLATION OF THE COVALENTLY
 ATTACHED BIOTIN IN THE FIRST STEP AND THE TRANSFER OF THE CARBOXYL
 GROUP TO PYRUVATE IN THE SECOND.
 CC -1- CATALYTIC ACTIVITY: ATP + pyruvate + HCO(3)(-) = ADP + phosphate +
 oxaloacetate.
 CC -1- COFACTOR: ATP, MAGNESIUM (OR MANGANESE OR COBALT), PYRUVATE AND
 BICARBONATE.
 CC -1- ENZYME REGULATION: INHIBITED BY MAGNESIUM, WHEN ITS CONCENTRATION
 EXCEEDED THE ATP ONE, AND BY HIGH CONCENTRATION OF ATP AND ALPHA-
 KETOGLUTARATE.
 CC -1- PATHWAY: GLUCONEOGENESIS.
 CC -1- SUBUNIT: HETEROCTAMER OF FOUR A AND FOUR B SUBUNITS.
 CC -1- MASS SPECTROMETRY: MW-64160; METHOD-MALDI.
 CC -1- MISCELLANEOUS: ITS OPTIMUM PH IS 8.5 AND THE OPTIMUM TEMPERATURE
 IS 80-90 DEGREES CELSIUS.
 CC -1- SIMILARITY: WITH OTHER BIOTIN CARBOXYLASES, LIPOAMIDE TRANSFERASES
 AND CARBAMYL PHOSPHATE SYNTHETASES.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL: U67563; AAB99233.1; -
 DR HSSP: P02905; 1A6X.
 DR TIGR: M01231; -
 DR InterPro: IPR001882; Biotin_attach.
 DR InterPro: IPR000089; Biotin_lipoyl.
 DR InterPro: IPR000891; HMG1-like.
 DR InterPro: IPR003379; Pyc-OADA.
 DR Pfam: PF00364; Biotin_lipoyl; 1.
 DR Pfam: PF00682; HMG1-like; 1.
 DR Pfam: PF02436; Pyc-OADA; 1.
 DR TIGRPFAMs: TIGR01108; Oada; 1.
 DR PROSITE: PS00188; BIOTIN; 1.
 KM Ligase: Multifunctional enzyme; Gluconeogenesis; Magnesium; Pyruvate;
 KM Biotin; Complete proteome.
 FT BINDING 52 60 PYRUVATE (BY SIMILARITY).
 FT BINDING 533 533 BIOTIN (BY SIMILARITY).
 SQ SEQUENCE 567 AA; 63907 MW; 5E07800622545628 CRC64;

Query Match 28.1%; Score 173; DB 1; Length 567;
 Best Local Similarity 50.7%; Pred. No. 2.2e-06;
 Matches 35; Conservative 12; Mismatches 22; Indels 0; Gaps 0;

QY 53 EGEIPAPLACTVSKILVKGDTVAKGVTVLEAKMKMETETINAPTDGKVEVLYKEDAV 112
 DB 498 EGAVTSPFRGMVTKIKKEDDKKKGVYVLEAKMKMETETINAPTDGKVEVLYKEDAV 557
 QY 113 QGGGGLIKI 121
 DB 558 NQGVDMIT 566

RESULT 5
 BCCP_STRMU STANDARD; PRT; 130 AA.
 AC P29337;
 DT 01-DEC-1992 (Rel. 24, Created)
 DT 01-DEC-1992 (Rel. 24, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE Biotin carboxyl carrier protein (BCCP).
 OS Streptococcus mutans.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_Taxid=1309;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=OT-041 / Serotype C;
 RX MEDLINE=93159778; PubMed=8431283;
 RA Wang D., Weyer M.M., Taricani M., Buckingham K., Sandham H.J.;
 RT "Biotin-containing protein as a cause of false positive clones in
 RT gene probing with streptavidin/biotin.";
 RL Biotechniques 14:209-212(1993).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL: M80523; AAA03702.1; -
 DR HSSP: P02905; 1A6X.
 DR InterPro: IPR001882; Biotin_attach.
 DR InterPro: IPR000891; Biotin_lipoyl.
 DR Pfam: PF00364; Biotin_lipoyl; 1.
 DR PROSITE: PS00188; BIOTIN; 1.
 KM Fatty acid biosynthesis; Biotin.

FT BINDING 96 96 BIOTIN (BY SIMILARITY).
 SQ SEQUENCE 130 AA; 13601 MW; A0D6025BC46FF00B CRC64;
 Query Match 27.3%; Score 168; DB 1; Length 130;
 Best Local Similarity 34.4%; Pred. No. 1.3e-06;
 Matches 45; Conservative 18; Mismatches 52; Indels 16; Gaps 3;

QY 2 KLVTVNGTAVDVVDVDRKSHENPMGTL--FGGGTGA-----PAPAAGAGAGR 50
 DB 4 KFKSIDGKEVLEMEISSESVPAATPIPTENTRAASDQKQSQSPAPATASAA-- 61
 QY 51 AGESEIPAPLACTVSKILVKGDTVAKGVTVLEAKMKMETETINAPTDGKVEVLYKED 110
 DB 62 ---NTMPAPMPGTILKLVVNGDTVSENPIMLEAKMKMETETINAGAGTVAIHVSSG 118
 QY 111 AVGGGGLIKI 121
 DB 119 TVDAGDMIT 129

RESULT 6
 BCCA_MYCLE STANDARD; PRT; 598 AA.
 ID BCCA_MYCLE
 AC P46392;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Acetyl-/propionyl-coenzyme A carboxylase alpha chain [includes: Biotin
 DE carboxylase (EC 6.3.4.14); Biotin carboxyl carrier protein (BCCP)].
 GN BCCA OR M0726 OR B1308_C1_129.
 OS Mycobacterium leprae.
 OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
 OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_Taxid=1769;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=94222829; PubMed=7909542;
 RA Norman E., de Smet K.A.L., Stoker N.G., Ratledge C., Wheeler P.R.,
 RA Dale J.W.;
 RT "Lipid synthesis in mycobacteria: characterization of the biotin
 RT carboxyl carrier protein genes from Mycobacterium leprae and M.
 RT tuberculosis".
 RL J. Bacteriol. 176:2525-2531(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Smith D.R., Robison K.;
 RL submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=TN;
 RX MEDLINE=21128732; PubMed=11234002;
 RA Cole S.T., Eigmeier K., Parkhill J., James K.D., Thomson N.R.,
 RA Wheeler P.R., Honore N., Garnier T., Churcher C., Harris D.,
 RA Mungall K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R.M., Devlin K., Duthoy S., Feltham T., Fraser A., Hamlin N.,
 RA Holtroyd S., Hornsby T., Jagels K., Lacroix C., Maclean J., Moule S.,
 RA Murphy L., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,
 RA Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squires R.,
 RA Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,
 RA Barrall B.G.;
 RT "Massive gene decay in the leprosy bacillus".
 RL Nature 409:1007-1011(2001).
 CC -1- FUNCTION: THIS PROTEIN CARRIES TWO FUNCTIONS: BIOTIN CARBOXYL
 CC CARRIER PROTEIN AND BIOTIN CARBOXYLTRANSFERASE.
 CC -1- CATALYTIC ACTIVITY: ATP + biotin-carboxyl-carrier protein + CO(2)
 CC = ADP + phosphate + carboxybiotin-carboxyl-carrier protein.
 CC -1- COFACTOR: BIOTIN.
 CC -1- PATHWAY: Long-chain fatty acid biosynthesis; first step.
 CC -1- SUBUNIT: MULTIMER COMPRISED OF 2 DIFFERENT SUBUNITS, THE LARGER
 CC ONE (63/64 kDa) HAS BIOTIN CARBOXYLASE AND BIOTIN CARRIER
 CC FUNCTIONS, WHILE THE SMALLER SUBUNIT POSSESSES CARBOXYLTRANSFERASE
 CC AND SUBSTRATE BINDING ACTIVITY.
 CC -1- SIMILARITY: TO OTHER BIOTIN-DEPENDENT ENZYMES AND CARBOXYL-

CC PHOSPHATE SYNTHETASES.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

DR EMBL: X63470; CAA45070.1; -

DR EMBL: U00012; AAA85920.1; -

DR EMBL: AL583919; CAC30235.1; -

DR HSSP: P24182; 1BNC.

DR Leptoma; ML0726; -

DR InterPro: IPR001882; Biotin_attach.

DR InterPro: IPR000089; Biotin_lipoyl.

DR InterPro: IPR000901; CPSase.

DR Pfam: PF00289; CPSase_L-chain; 1.

DR Pfam: PF00364; biotin_lipoyl; 1.

DR Pfam: PF02785; biotin_card_C; 1.

DR Pfam: PF02786; CPSase_L-D2; 1.

DR PRINTS: PR00098; CPSASE.

DR PROSITE: PS00188; BIOTIN; 1.

DR PROSITE: PS00867; CPSASE_1; 1.

DR PROSITE: PS00867; CPSASE_2; 1.

DR Fatty acid biosynthesis; Ligase; Biotin; Multifunctional enzyme;

KW ATP-binding; Complete proteome.

FT DOMAIN 11 441 BIOTIN CARBOXYLASE.

FT NP_BIND 532 598 BIOTIN CARBOXYL CARRIER PROTEIN.

FT ACT_SITE 169 174 ATP (BY SIMILARITY).

FT BINDING 299 299 BY SIMILARITY.

FT BINDING 564 564 BIOTIN (BY SIMILARITY).

FT CONFLICT 30 30 D -> H (IN REF. 1).

FT CONFLICT 30 30 D -> H (IN REF. 1).

SO SEQUENCE 598 AA; 63863 MW; 5E2E291D7C54515D CRC64;

Query Match 26.5%; Score 163; DB 1; Length 598;

Best Local Similarity 37.3%; Pred. No. 1.4e-05;

Matches 44; Conservative 15; Mismatches 57; Indels 2; Gaps 2;

QY 5 VTVNGTAVDVVDKSHENPMGTLFGGTGAPAPAGAGAGAGAGE-IPAPLAGT 63

DB 481 VEVGSRLEVLDPDLALANPAGCNP-AGVIRKKRPKRGRGHGCAIAGSDAVTAPMGCT 539

QY 64 VSKILVNEGTVKAGQIVLVLEAMKETEINAPDGVKEKVLKERDAVGGGGLIKI 121

DB 540 VVKVAVAGQVTMGDVLVYVLEAMKEMENPVTAHKDGIITGLVAGTAITGTVLAEI 597

RESULT 7

MCCA_MOUSE STANDARD; PRT; 717 AA.

AC 099MR8; Q9D8R2;

DT 15-JUN-2002 (Rel. 41, Created)

DT 15-JUN-2002 (Rel. 41, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Methylcrotonyl-CoA carboxylase alpha chain, mitochondrial precursor

DE (EC 6.4.1.4) (3-methylcrotonyl-CoA carboxylase 1) (MCCase alpha

DE subunit) (3-methylcrotonyl-CoA:carbon dioxide ligase alpha subunit).

GN MCCCL OR MCCA.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6;

RX MEDLINE=21102410; PubMed=11181649;

RA Baumgartner M.R., Almashanu S., Suormala T., Obie C., Cole R.N.,

RA Packman S., Baumgartner E.R., Valle D.;

RT "The molecular basis of human 3-methylcrotonyl-CoA carboxylase

RT deficiency.";

RL J. Clin. Invest. 107:495-504(2001).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Pancreas;

RX MEDLINE=21085660; PubMed=11217851;

RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,

RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,

RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K.I.,

RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,

RA Kadota K., Matsuda H.A., Ashburner M., Batilov S., Casavant T.,

RA Flisbachmann W., Gaasterland T., Gissi C., King B., Kochava H.,

RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,

RA Schriml L.M., Straubli F., Suzuki R., Tomita M., Wagner L., Washio T.,

RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,

RA Blake J., Boilelli D., Bojunga N., Carninci P., de Bonaldo M.F.,

RA Brownstein M.J., Bull C., Fletcher C., Fujita M., Gariboldi M.,

RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kanlaya M., Lee N.H.,

RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,

RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,

RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,

RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,

RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawai J., Kohsaki S.,

RA Hayashizaki Y.;

RT "Functional annotation of a full-length mouse cDNA collection.";

RL Nature 409:685-690(2001).

RN [3]

RP SEQUENCE FROM N.A.

RC TISSUE=Kidney;

RA Strausberg R.;

RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.

CC -1- CATALYTIC ACTIVITY: ATP + 3-methylcrotonyl-CoA + HCO(3)(-) - ADP +

CC phosphate + 3-methylglutacetyl-CoA.

CC -1- Cofactor: Biotin.

CC -1- PATHWAY: leucine catabolism.

CC -1- SUBUNIT: Probably a dodecamer composed of six biotin-containing

CC alpha subunits and six beta subunits.

CC -1- SUBCELLULAR LOCATION: Mitochondrial matrix (By similarity).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

DR EMBL: AF310338; AAC50244.1; -

DR EMBL: AK007782; BAB25253.1; -

DR EMBL: BC021382; AAH21382.1; -

DR HSSP: P24182; 1DVL.

DR MGD; MGI:1919289; Mcccl.

DR InterPro: IPR001882; Biotin_attach.

DR InterPro: IPR000089; Biotin_lipoyl.

DR InterPro: IPR000901; CPSase.

DR Pfam: PF00289; CPSase_L-chain; 1.

DR Pfam: PF00364; biotin_lipoyl; 1.

DR Pfam: PF02785; biotin_card_C; 1.

DR Pfam: PF02786; CPSase_L-D2; 1.

DR PROSITE: PS00188; BIOTIN; 1.

DR PROSITE: PS00867; CPSASE_1; 1.

DR PROSITE: PS00867; CPSASE_2; 1.

KW Mitochondrion; Ligase; Biotin; ATP-binding; Transit peptide.

FT TRANSIT 1 44

FT CHAIN 45 717 MITOCHONDRIAL COA CARBOXYLASE ALPHA

FT NP_BIND 205 210 CHAIN.

FT ACT_SITE 335 335 ATP (POTENTIAL).

FT BINDING 677 677 BY SIMILARITY.

FT DOMAIN 709 714 BIOTIN (BY SIMILARITY).

FT CONFLICT 324 324 POLY-GLU.

FT CONFLICT 324 324 R -> K (IN REF. 3).

FT CONFLICT 507 507 A -> P (IN REF. 1).

SO SEQUENCE 717 AA; 79343 MW; F653FE7AC1E5A90 CRC64;

Query Match 25.9%; Score 159; DB 1; Length 717;

Best Local Similarity 36.0%; Pred. No. 3.3e-05;

Matches 40; Conservative 17; Mismatches 42; Indels 12; Gaps 3;

QY 3 LKVTNGTAVDQVDVSKH---ENPMGTLIFGCGTG-GAPAPAGAGAGKAGEEIP 57
 DB 594 LKSSINGVA-----RKSKEILLDNTVHLFSMEGSLIEVGPVPRKYLSPVSAEGAGGTI 646
 QY 58 APLAGTYSKILVREGDVKAGQTVLVEAMKMEFEINAPTEGKVKYVKE 108
 DB 647 APMGTIEKVFVKAGDRKADSLMVMIMKMEHTIKAPKDKIRIKVFSE 697

RESULT 8
 MCCA_HUMAN
 ID MCCA_HUMAN STANDARD; PRT; 725 AA.
 AC Q96R03; Q9H959; Q9NS97;
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Methylcrotonyl-CoA carboxylase alpha chain, mitochondrial precursor
 DE (EC 6.4.1.4) (3-methylcrotonyl-CoA carboxylase 1) (MCCase alpha
 subunit) (3-methylcrotonyl-CoA:carbon dioxide ligase alpha subunit).
 GN MCCC1 OR MCCA.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A., AND VARIANTS MCGI ARG-325 AND SER-385;
 RX PubMed=11170888;
 RA Gallardo M.E., Desviat L.R., Rodriguez J.M., Esparza-Gordillo J.,
 RA Perez-Cerdá C., Perez B., Rodriguez-Pombo P., Criado O., Sanz R.,
 RA Morton D.H., Gibson K.M., Le T.P., Ribes A., Rodriguez de Cordoba S.,
 RA Ugarde M., Penalva M.A.;
 RT "The molecular basis of 3-methylcrotonylglycinuria, a disorder of
 RT leucine catabolism";
 RL Am. J. Hum. Genet. 68:334-346(2001).
 RN [2]
 RP SEQUENCE FROM N.A., AND VARIANT HIS-464.
 RX MEDLINE=21295033; PubMed=11401427;
 RA Odata K., Fukuda T., Morishita R., Abe S., Asakawa S., Yamaguchi S.,
 RA Yoshino M., Ihara K., Murayama K., Shimizu N., Kondo I.;
 RT "Human biotin-containing subunit of 3-methylcrotonyl-CoA carboxylase
 RT gene (MCCA): cDNA sequence, genomic organization, localization to
 RT chromosome band 3q27, and expression";
 RL Genomics 72:145-152(2001).
 RN [3]
 RP SEQUENCE FROM N.A., AND VARIANT MCGI PHE-535.
 RX MEDLINE=21299419; PubMed=11406611;
 RA Holzinger A., Roeschinger W., Lagler F., Mayerhofer P.U., Lichtner P.,
 RA Kattenfeld T., Thuy L.P., Nyhan W.L., Koch H.G., Muntau A.C.,
 RA Roscher A.A.;
 RT "Cloning of the human MCCA and MCB genes and mutations therein reveal
 RT the molecular cause of 3-methylcrotonyl-CoA: carboxylase
 RT deficiency";
 RL Hum. Mol. Genet. 10:1299-1306(2001).
 RN [4]
 RP SEQUENCE FROM N.A., AND VARIANTS MCGI VAL-289; SER-385; PRO-437 AND
 RP HIS-532.
 RX PubMed=11181649;
 RA Baumgartner M.R., Almashanu S., Suormala T., Obie C., Cole R.N.,
 RA Packman S., Baumgartner E.R., Valle D.;
 RT "The molecular basis of human 3-methylcrotonyl-CoA carboxylase
 RT deficiency";
 RL J. Clin. Invest. 107:495-504(2001).
 RN [5]
 RP SEQUENCE FROM N.A.
 RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
 RA Nishikawa T., Nagai K., Sugano S., Aotsuka S., Yoshikawa Y.,
 RA Matsunawa H., Ishii S., Kawai Y., Saito K., Yamamoto J., Wakamatsu A.,
 RA Nakamura Y., Nagahari K., Masuko Y., Sasaki N.;
 RT "NEO human cDNA sequencing project";
 RL Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.
 RN [6]

RP SEQUENCE FROM N.A.
 RC TISSUE=Skeletal muscle;
 RA Strausberg R.;
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: ATP + 3-methylcrotonyl-CoA + H₂O(3)(-) - ADP +
 CC phosphate + 3-methylglutacetyl-CoA;
 CC -1- COFACTOR: Biotin.
 CC -1- PATHWAY: Leucine catabolism.
 CC -1- SUBUNIT: Probably a dodecamer composed of six biotin-containing
 CC alpha subunits and six beta subunits.
 CC -1- SUBCELLULAR LOCATION: Mitochondrial matrix.
 CC -1- DISEASE: Defects in MCCC1 are the cause of 3-
 CC methylcrotonylglycinuria type I (MCGI, CGA or CG2). MCGI is a
 CC recessive disease that is characterized by muscular hypotonia and
 CC atrophy, probably of spinal origin.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AF310972; AAC53095.1; -
 DR EMBL: AB029826; BAA9407.1; -
 DR EMBL: AF297332; AAK67986.1; -
 DR EMBL: AF310339; AAG50245.1; -
 DR EMBL: AK023051; BAB14377.1; -
 DR EMBL: BC004214; AAH04214.1; -
 DR EMBL: BC004187; AAH04187.1; -
 DR Genew: HGNC:6936; MCCC1.
 DR MIM: 210200; -
 DR InterPro: IPR001882; Biotin_attach.
 DR InterPro: IPR000089; Biotin_lipoyl.
 DR InterPro: IPR000901; CPase.
 DR Pfam: PF00289; CPase_L_chain; 1.
 DR Pfam: PF00364; Biotin_lipoyl; 1.
 DR Pfam: PF02785; Biotin_carb_C; 1.
 DR Pfam: PF02786; CPase_L_D2; 1.
 DR PROSITE: PS00188; Biotin; 1.
 DR PROSITE: PS00867; CPASE_2; 1.
 KW Mitochondrion; Ligase; Biotin; ATP-binding; Transit peptide;
 KW Disease mutation; Polymorphism.
 FT TRANSIT 1 47
 FT CHAIN 48 725
 FT FT
 FT NP_BIND 209 214
 FT ACT_SITE 339 339
 FT BINDING 681 681
 FT DOMAIN 538 541
 FT DOMAIN 713 718
 FT VARIANT 289 289
 FT VARIANT 325 325
 FT VARIANT 385 385
 FT VARIANT 437 437
 FT VARIANT 464 464
 FT VARIANT 532 532
 FT VARIANT 535 535
 FT VARIANT 549 549
 FT VARIANT 553 553
 FT VARIANT 555 555
 FT VARIANT 557 557
 FT VARIANT 559 559
 FT VARIANT 561 561
 FT VARIANT 563 563
 FT VARIANT 565 565
 FT VARIANT 567 567
 FT VARIANT 569 569
 FT VARIANT 571 571
 FT VARIANT 573 573
 FT VARIANT 575 575
 FT VARIANT 577 577
 FT VARIANT 579 579
 FT VARIANT 581 581
 FT VARIANT 583 583
 FT VARIANT 585 585
 FT VARIANT 587 587
 FT VARIANT 589 589
 FT VARIANT 591 591
 FT VARIANT 593 593
 FT VARIANT 595 595
 FT VARIANT 597 597
 FT VARIANT 599 599
 FT VARIANT 601 601
 FT VARIANT 603 603
 FT VARIANT 605 605
 FT VARIANT 607 607
 FT VARIANT 609 609
 FT VARIANT 611 611
 FT VARIANT 613 613
 FT VARIANT 615 615
 FT VARIANT 617 617
 FT VARIANT 619 619
 FT VARIANT 621 621
 FT VARIANT 623 623
 FT VARIANT 625 625
 FT VARIANT 627 627
 FT VARIANT 629 629
 FT VARIANT 631 631
 FT VARIANT 633 633
 FT VARIANT 635 635
 FT VARIANT 637 637
 FT VARIANT 639 639
 FT VARIANT 641 641
 FT VARIANT 643 643
 FT VARIANT 645 645
 FT VARIANT 647 647
 FT VARIANT 649 649
 FT VARIANT 651 651
 FT VARIANT 653 653
 FT VARIANT 655 655
 FT VARIANT 657 657
 FT VARIANT 659 659
 FT VARIANT 661 661
 FT VARIANT 663 663
 FT VARIANT 665 665
 FT VARIANT 667 667
 FT VARIANT 669 669
 FT VARIANT 671 671
 FT VARIANT 673 673
 FT VARIANT 675 675
 FT VARIANT 677 677
 FT VARIANT 679 679
 FT VARIANT 681 681
 FT VARIANT 683 683
 FT VARIANT 685 685
 FT VARIANT 687 687
 FT VARIANT 689 689
 FT VARIANT 691 691
 FT VARIANT 693 693
 FT VARIANT 695 695
 FT VARIANT 697 697
 FT VARIANT 699 699
 FT VARIANT 701 701
 FT VARIANT 703 703
 FT VARIANT 705 705
 FT VARIANT 707 707
 FT VARIANT 709 709
 FT VARIANT 711 711
 FT VARIANT 713 713
 FT VARIANT 715 715
 FT VARIANT 717 717
 FT VARIANT 719 719
 FT VARIANT 721 721
 FT VARIANT 723 723
 FT VARIANT 725 725
 FT VARIANT 727 727
 FT VARIANT 729 729
 FT VARIANT 731 731
 FT VARIANT 733 733
 FT VARIANT 735 735
 FT VARIANT 737 737
 FT VARIANT 739 739
 FT VARIANT 741 741
 FT VARIANT 743 743
 FT VARIANT 745 745
 FT VARIANT 747 747
 FT VARIANT 749 749
 FT VARIANT 751 751
 FT VARIANT 753 753
 FT VARIANT 755 755
 FT VARIANT 757 757
 FT VARIANT 759 759
 FT VARIANT 761 761
 FT VARIANT 763 763
 FT VARIANT 765 765
 FT VARIANT 767 767
 FT VARIANT 769 769
 FT VARIANT 771 771
 FT VARIANT 773 773
 FT VARIANT 775 775
 FT VARIANT 777 777
 FT VARIANT 779 779
 FT VARIANT 781 781
 FT VARIANT 783 783
 FT VARIANT 785 785
 FT VARIANT 787 787
 FT VARIANT 789 789
 FT VARIANT 791 791
 FT VARIANT 793 793
 FT VARIANT 795 795
 FT VARIANT 797 797
 FT VARIANT 799 799
 FT VARIANT 801 801
 FT VARIANT 803 803
 FT VARIANT 805 805
 FT VARIANT 807 807
 FT VARIANT 809 809
 FT VARIANT 811 811
 FT VARIANT 813 813
 FT VARIANT 815 815
 FT VARIANT 817 817
 FT VARIANT 819 819
 FT VARIANT 821 821
 FT VARIANT 823 823
 FT VARIANT 825 825
 FT VARIANT 827 827
 FT VARIANT 829 829
 FT VARIANT 831 831
 FT VARIANT 833 833
 FT VARIANT 835 835
 FT VARIANT 837 837
 FT VARIANT 839 839
 FT VARIANT 841 841
 FT VARIANT 843 843
 FT VARIANT 845 845
 FT VARIANT 847 847
 FT VARIANT 849 849
 FT VARIANT 851 851
 FT VARIANT 853 853
 FT VARIANT 855 855
 FT VARIANT 857 857
 FT VARIANT 859 859
 FT VARIANT 861 861
 FT VARIANT 863 863
 FT VARIANT 865 865
 FT VARIANT 867 867
 FT VARIANT 869 869
 FT VARIANT 871 871
 FT VARIANT 873 873
 FT VARIANT 875 875
 FT VARIANT 877 877
 FT VARIANT 879 879
 FT VARIANT 881 881
 FT VARIANT 883 883
 FT VARIANT 885 885
 FT VARIANT 887 887
 FT VARIANT 889 889
 FT VARIANT 891 891
 FT VARIANT 893 893
 FT VARIANT 895 895
 FT VARIANT 897 897
 FT VARIANT 899 899
 FT VARIANT 901 901
 FT VARIANT 903 903
 FT VARIANT 905 905
 FT VARIANT 907 907
 FT VARIANT 909 909
 FT VARIANT 911 911
 FT VARIANT 913 913
 FT VARIANT 915 915
 FT VARIANT 917 917
 FT VARIANT 919 919
 FT VARIANT 921 921
 FT VARIANT 923 923
 FT VARIANT 925 925
 FT VARIANT 927 927
 FT VARIANT 929 929
 FT VARIANT 931 931
 FT VARIANT 933 933
 FT VARIANT 935 935
 FT VARIANT 937 937
 FT VARIANT 939 939
 FT VARIANT 941 941
 FT VARIANT 943 943
 FT VARIANT 945 945
 FT VARIANT 947 947
 FT VARIANT 949 949
 FT VARIANT 951 951
 FT VARIANT 953 953
 FT VARIANT 955 955
 FT VARIANT 957 957
 FT VARIANT 959 959
 FT VARIANT 961 961
 FT VARIANT 963 963
 FT VARIANT 965 965
 FT VARIANT 967 967
 FT VARIANT 969 969
 FT VARIANT 971 971
 FT VARIANT 973 973
 FT VARIANT 975 975
 FT VARIANT 977 977
 FT VARIANT 979 979
 FT VARIANT 981 981
 FT VARIANT 983 983
 FT VARIANT 985 985
 FT VARIANT 987 987
 FT VARIANT 989 989
 FT VARIANT 991 991
 FT VARIANT 993 993
 FT VARIANT 995 995
 FT VARIANT 997 997
 FT VARIANT 999 999
 FT VARIANT 1001 1001
 FT VARIANT 1003 1003
 FT VARIANT 1005 1005
 FT VARIANT 1007 1007
 FT VARIANT 1009 1009
 FT VARIANT 1011 1011
 FT VARIANT 1013 1013
 FT VARIANT 1015 1015
 FT VARIANT 1017 1017
 FT VARIANT 1019 1019
 FT VARIANT 1021 1021
 FT VARIANT 1023 1023
 FT VARIANT 1025 1025
 FT VARIANT 1027 1027
 FT VARIANT 1029 1029
 FT VARIANT 1031 1031
 FT VARIANT 1033 1033
 FT VARIANT 1035 1035
 FT VARIANT 1037 1037
 FT VARIANT 1039 1039
 FT VARIANT 1041 1041
 FT VARIANT 1043 1043
 FT VARIANT 1045 1045
 FT VARIANT 1047 1047
 FT VARIANT 1049 1049
 FT VARIANT 1051 1051
 FT VARIANT 1053 1053
 FT VARIANT 1055 1055
 FT VARIANT 1057 1057
 FT VARIANT 1059 1059
 FT VARIANT 1061 1061
 FT VARIANT 1063 1063
 FT VARIANT 1065 1065
 FT VARIANT 1067 1067
 FT VARIANT 1069 1069
 FT VARIANT 1071 1071
 FT VARIANT 1073 1073
 FT VARIANT 1075 1075
 FT VARIANT 1077 1077
 FT VARIANT 1079 1079
 FT VARIANT 1081 1081
 FT VARIANT 1083 1083
 FT VARIANT 1085 1085
 FT VARIANT 1087 1087
 FT VARIANT 1089 1089
 FT VARIANT 1091 1091
 FT VARIANT 1093 1093
 FT VARIANT 1095 1095
 FT VARIANT 1097 1097
 FT VARIANT 1099 1099
 FT VARIANT 1101 1101
 FT VARIANT 1103 1103
 FT VARIANT 1105 1105
 FT VARIANT 1107 1107
 FT VARIANT 1109 1109
 FT VARIANT 1111 1111
 FT VARIANT 1113 1113
 FT VARIANT 1115 1115
 FT VARIANT 1117 1117
 FT VARIANT 1119 1119
 FT VARIANT 1121 1121
 FT VARIANT 1123 1123
 FT VARIANT 1125 1125
 FT VARIANT 1127 1127
 FT VARIANT 1129 1129
 FT VARIANT 1131 1131
 FT VARIANT 1133 1133
 FT VARIANT 1135 1135
 FT VARIANT 1137 1137
 FT VARIANT 1139 1139
 FT VARIANT 1141 1141
 FT VARIANT 1143 1143
 FT VARIANT 1145 1145
 FT VARIANT 1147 1147
 FT VARIANT 1149 1149
 FT VARIANT 1151 1151
 FT VARIANT 1153 1153
 FT VARIANT 1155 1155
 FT VARIANT 1157 1157
 FT VARIANT 1159 1159
 FT VARIANT 1161 1161
 FT VARIANT 1163 1163
 FT VARIANT 1165 1165
 FT VARIANT 1167 1167
 FT VARIANT 1169 1169
 FT VARIANT 1171 1171
 FT VARIANT 1173 1173
 FT VARIANT 1175 1175
 FT VARIANT 1177 1177
 FT VARIANT 1179 1179
 FT VARIANT 1181 1181
 FT VARIANT 1183 1183
 FT VARIANT 1185 1185
 FT VARIANT 1187 1187
 FT VARIANT 1189 1189
 FT VARIANT 1191 1191
 FT VARIANT 1193 1193
 FT VARIANT 1195 1195
 FT VARIANT 1197 1197
 FT VARIANT 1199 1199
 FT VARIANT 1201 1201
 FT VARIANT 1203 1203
 FT VARIANT 1205 1205
 FT VARIANT 1207 1207
 FT VARIANT 1209 1209
 FT VARIANT 1211 1211
 FT VARIANT 1213 1213
 FT VARIANT 1215 1215
 FT VARIANT 1217 1217
 FT VARIANT 1219 1219
 FT VARIANT 1221 1221
 FT VARIANT 1223 1223
 FT VARIANT 1225 1225
 FT VARIANT 1227 1227
 FT VARIANT 1229 1229
 FT VARIANT 1231 1231
 FT VARIANT 1233 1233
 FT VARIANT 1235 1235
 FT VARIANT 1237 1237
 FT VARIANT 1239 1239
 FT VARIANT 1241 1241
 FT VARIANT 1243 1243
 FT VARIANT 1245 1245
 FT VARIANT 1247 1247
 FT VARIANT 1249 1249
 FT VARIANT 1251 1251
 FT VARIANT 1253 1253
 FT VARIANT 1255 1255
 FT VARIANT 1257 1257
 FT VARIANT 1259 1259
 FT VARIANT 1261 1261
 FT VARIANT 1263 1263
 FT VARIANT 1265 1265
 FT VARIANT 1267 1267
 FT VARIANT 1269 1269
 FT VARIANT 1271 1271
 FT VARIANT 1273 1273
 FT VARIANT 1275 1275
 FT VARIANT 1277 1277
 FT VARIANT 1279 1279
 FT VARIANT 1281 1281
 FT VARIANT 1283 1283
 FT VARIANT 1285 1285
 FT VARIANT 1287 1287
 FT VARIANT 1289 1289
 FT VARIANT 1291 1291
 FT VARIANT 1293 1293
 FT VARIANT 1295 1295
 FT VARIANT 1297 1297
 FT VARIANT 1299 1299
 FT VARIANT 1301 1301
 FT VARIANT 1303 1303
 FT VARIANT 1305 1305
 FT VARIANT 1307 1307
 FT VARIANT 1309 1309
 FT VARIANT 1311 1311
 FT VARIANT 1313 1313
 FT VARIANT 1315 1315
 FT VARIANT 1317 1317
 FT VARIANT 1319 1319
 FT VARIANT 1321 1321
 FT VARIANT 1323 1323
 FT VARIANT 1325 1325
 FT VARIANT 1327 1327
 FT VARIANT 1329 1329
 FT VARIANT 1331 1331
 FT VARIANT 1333 1333
 FT VARIANT 1335 1335
 FT VARIANT 1337 1337
 FT VARIANT 1339 1339
 FT VARIANT 1341 1341
 FT VARIANT 1343 1343
 FT VARIANT 1345 1345
 FT VARIANT 1347 1347
 FT VARIANT 1349 1349
 FT VARIANT 1351 1351
 FT VARIANT 1353 1353
 FT VARIANT 1355 1355
 FT VARIANT 1357 1357
 FT VARIANT 1359 1359
 FT VARIANT 1361 1361
 FT VARIANT 1363 1363
 FT VARIANT 1365 1365
 FT VARIANT 1367 1367
 FT VARIANT 1369 1369
 FT VARIANT 1371 1371
 FT VARIANT 1373 1373
 FT VARIANT 1375 1375
 FT VARIANT 1377 1377
 FT VARIANT 1379 1379
 FT VARIANT 1381 1381
 FT VARIANT 1383 1383
 FT VARIANT 1385 1385
 FT VARIANT 1387 1387
 FT VARIANT 1389 1389
 FT VARIANT 1391 1391
 FT VARIANT 1393 1393
 FT VARIANT 1395 1395
 FT VARIANT 1397 1397
 FT VARIANT 1399 1399
 FT VARIANT 1401 1401
 FT VARIANT 1403 1403
 FT VARIANT 1405 1405
 FT VARIANT 1407 1407
 FT VARIANT 1409 1409
 FT VARIANT 1411 1411
 FT VARIANT 1413 1413
 FT VARIANT 1415 1415
 FT VARIANT 1417 1417
 FT VARIANT 1419 1419
 FT VARIANT 1421 1421
 FT VARIANT 1423 1423
 FT VARIANT 1425 1425
 FT VARIANT 1427 1427
 FT VARIANT 1429 1429
 FT VARIANT 1431 1431
 FT VARIANT 1433 1433
 FT VARIANT 1435 1435
 FT VARIANT 1437 1437
 FT VARIANT 1439 1439
 FT VARIANT 1441 1441
 FT VARIANT 1443 1443
 FT VARIANT 1445 1445
 FT VARIANT 1447 1447
 FT VARIANT 1449 1449
 FT VARIANT 1451 1451
 FT VARIANT 1453 1453
 FT VARIANT 1455 1455
 FT VARIANT 1457 1457
 FT VARIANT 1459 1459
 FT VARIANT 1461 1461
 FT VARIANT 1463 1463
 FT VARIANT 1465 1465
 FT VARIANT 1467 1467
 FT VARIANT 1469 1469
 FT VARIANT 1471 1471
 FT VARIANT 1473 1473
 FT VARIANT 1475 1475
 FT VARIANT 1477 1477
 FT VARIANT 1479 1479
 FT VARIANT 1481 1481
 FT VARIANT 1483 1483
 FT VARIANT 1485 1485
 FT VARIANT 1487 1487
 FT VARIANT 1489 1489
 FT VARIANT 1491 1491
 FT VARIANT 1493 1493
 FT VARIANT 1495 1495
 FT VARIANT 1497 1497
 FT VARIANT 1499 1499
 FT VARIANT 1501 1501
 FT VARIANT 1503 1503
 FT VARIANT 1505 1505
 FT VARIANT 1507 1507
 FT VARIANT 1509 1509
 FT VARIANT 1511 1511
 FT VARIANT 1513 1513
 FT VARIANT 1515 1515
 FT VARIANT 1517 1517
 FT VARIANT 1519 1519
 FT VARIANT 1521 1521
 FT VARIANT 1523 1523
 FT VARIANT 1525 1525
 FT VARIANT 1527 1527
 FT VARIANT 1529 1529
 FT VARIANT 1531 1531
 FT VARIANT 1533 1533
 FT VARIANT 1535 1535
 FT VARIANT 1537 1537
 FT VARIANT 1539 1539
 FT VARIANT 1541 1541
 FT VARIANT 1543 1543
 FT VARIANT 1545 1545
 FT VARIANT 1547 1547
 FT VARIANT 1549 1549
 FT VARIANT 1551 1551
 FT VARIANT 1553 1553
 FT VARIANT 1555 1555
 FT VARIANT 1557 1557
 FT VARIANT 1559 1559
 FT VARIANT 1561 1561
 FT VARIANT 1563 1563
 FT VARIANT 1565 1565
 FT VARIANT 1567 1567
 FT VARIANT 1569 1569
 FT VARIANT 1571 1571
 FT VARIANT 1573 1573
 FT VARIANT 1575 1575
 FT VARIANT 1577 1577
 FT VARIANT 1579 1579
 FT VARIANT 1581 1581
 FT VARIANT 1583 1583
 FT VARIANT 1585 1585
 FT VARIANT 1587 1587
 FT VARIANT 1589 1589
 FT VARIANT 1591 1591
 FT VARIANT 1593 1593
 FT VARIANT 1595 1595
 FT VARIANT 1597 1597
 FT VARIANT 1599 1599
 FT VARIANT 1601 1601
 FT VARIANT 1603 1603
 FT VARIANT 1605 1605
 FT VARIANT 1607 1607
 FT VARIANT 1609 1609
 FT VARIANT 1611 1611
 FT VARIANT 1613 1613
 FT VARIANT 1615 1615
 FT VARIANT 1617 1617
 FT VARIANT 1619 1619
 FT VARIANT 1621 1621
 FT VARIANT 1623 1623
 FT VARIANT 1625 1625
 FT VARIANT 1627 1627
 FT VARIANT 1629 1629
 FT VARIANT 1631 1631
 FT VARIANT 1633 1633
 FT VARIANT 1635 1635
 FT VARIANT 1637 1637
 FT VARIANT 1639 1639
 FT VARIANT 1641 1641
 FT VARIANT 1643 1643
 FT VARIANT 1645 1645
 FT VARIANT 1647 1647
 FT VARIANT 1649 1649
 FT VARIANT 1651 1651
 FT VARIANT 1653 1653
 FT VARIANT 1655 1655
 FT VARIANT 1657 1657
 FT VARIANT 1659 1659
 FT VARIANT 1661 1661
 FT VARIANT 1663 1663
 FT VARIANT 1665 1665
 FT VARIANT 1667 1667
 FT VARIANT 1669 1669
 FT VARIANT 1671 1671
 FT VARIANT 1673 1673
 FT VARIANT 1675 1675
 FT VARIANT 1677 1677
 FT VARIANT 1679 1679
 FT VARIANT 1681 1681
 FT VARIANT 1683 1683
 FT VARIANT 1685 1685
 FT VARIANT 1687 1687
 FT VARIANT 1689 1689
 FT VARIANT 1691 1691
 FT VARIANT 1693 1693
 FT VARIANT 1695 1695
 FT VARIANT 1697 1697
 FT VARIANT 1699 1699
 FT VARIANT 1701 1701
 FT VARIANT 1703 1703
 FT VARIANT 1705 1705
 FT VARIANT 1707 1707
 FT VARIANT 1709 1709
 FT VARIANT 1711 1711
 FT VARIANT 1713 1713
 FT VARIANT 1715 1715
 FT VARIANT 1717 1717
 FT VARIANT 1719 1719
 FT VARIANT 1721 1721
 FT VARIANT 1723 1723
 FT VARIANT 1725 1725
 FT VARIANT 1727 1727
 FT VARIANT 1729 1729
 FT VARIANT 1731 1731
 FT VARIANT 1733 1733
 FT VARIANT 1735 1735
 FT VARIANT 1737 1737
 FT VARIANT 1739 1739
 FT VARIANT 1741 1741
 FT VARIANT 1743 1743
 FT VARIANT 1745 1745
 FT VARIANT 1747 1747
 FT VARIANT 1749 1749
 FT VARIANT 1751 1751
 FT VARIANT 1753 1753
 FT VARIANT 1755 1755
 FT VARIANT 1757 1757
 FT VARIANT 1759 1759
 FT VARIANT 1761 1761
 FT VARIANT 1763 1763
 FT VARIANT 1765 1765
 FT VARIANT 1767 1767
 FT VARIANT 1769 1769
 FT VARIANT 1771 1771
 FT VARIANT 1773 1773
 FT VARIANT 1775 1775
 FT VARIANT 1777 1777
 FT VARIANT 1779 1779
 FT VARIANT 1781 1781
 FT VARIANT 1783 1783
 FT VARIANT 1785 1785
 FT VARIANT 1787 1787
 FT VARIANT 1789 1789
 FT VARIANT 1791 1791
 FT VARIANT 1793 1793
 FT VARIANT 1795 1795

OY 3 LKVTNGTAVDVVDVDSKSHENPMGTILFGGCTGG-----APAPAAGAGAGAGGGEIIPA 58
 DB 598 LKCSVNGVASKAKLII---LEN---TILFSEKSGEIEIDIPKYLISVSSQETGGPLA 651
 OY 59 PLAGVSVLTKKEDPTKAGQTVLVLKMKMETEINAPTDGVEVYKKE 108
 DB 652 PMGTIEKVEFKAGDKVAKGDSLMVIMKMKMETHTKSPDGTAKVYRE 701

RESULT 9
 BCCP_LYCES STANDARD: PRT: 70 AA.
 AC POS115:
 DT 13-AUG-1987 (Rel. 05, Created)
 DT 13-AUG-1987 (Rel. 05, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Biotin carboxyl carrier protein of acetyl-CoA carboxylase (BCCP)
 DE (Fragment).
 OS Lycopersicon esculentum (Tomato).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 CC Asteridae; eusterids I; Solanales; Solanaceae; Solanum.
 CC NCBL_TaxID=4081;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE-87231088; PubMed-3588314;
 RA Hoffman N.E., Pichersky E., Cashmore A.R.;
 RT "A tomato cDNA encoding a biotin-binding protein.";
 RL Nucleic Acids Res. 15:3928-3928(1987).
 CC -1- FUNCTION: THIS PROTEIN IS A COMPONENT OF THE ACETYL COENZYME A
 CARBOXYLASE COMPLEX; FIRST, BIOTIN CARBOXYLASE CATALYZES THE
 TRANSFERS THE CARBOXYL GROUP TO FORM MALONYL-COA.
 CC -1- PATHWAY: Long-chain fatty acid biosynthesis; first step.
 CC -1- SUBCELLULAR LOCATION: Chloroplast.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@sib.ch).
 CC -----
 CC DR EMBL; Y00144; CA68339.1; -
 DR PIR; A29271; A29271.
 DR HSSP; P02905; 1BDO.
 DR InterPro: IPR001882; Biotin_attach.
 DR InterPro: IPR000089; Biotin_1lpoyl.
 DR Pfam; PF00364; Biotin_1lpoyl; 1.
 DR PROSITE; PS00188; BIOTIN; 1.
 KW Fatty acid biosynthesis; Biotin; Chloroplast.
 FT NON_TER 1
 FT BINDING 35
 FT SEQUENCE 70 AA; 7416 MW; D75D018C8BD016BC CRC64;

Query Match 23.3%; Score 143; DB 1; Length 70;
 Best Local Similarity 48.4%; Pred. No. 5.8e-05;
 Matches 30; Conservative 11; Mismatches 21; Indels 0; Gaps 0;

OY 54 GEIRPLPLAGTYSKIYKEDDYKAGQTVLVLKMKMETEINAPTDGKVEYKVEDAVO 113
 DB 1 GTTVAPMVGLEVKVLKDGKEVQEPVLYLEAMKMEHVAKAPANGYSGLEIKGVSVO 60

OY 114 GG 115
 DB 61 DG 62

RESULT 10
 MCCA_ARATH STANDARD: PRT: 734 AA.
 AC Q42523; Q9SA61;

DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Methylcrotonyl-CoA carboxylase alpha chain, mitochondrial precursor
 DE subunit) (3-methylcrotonyl-CoA:carbon dioxolase 1) (MCCase alpha
 DE subunit) (3-methylcrotonyl-CoA carboxylase 1) (MCCase alpha
 DE subunit) (3-methylcrotonyl-CoA carboxylase 1) (MCCase alpha
 GN MCCA OR AT1G03090 OR F1003.8 OR F1003.9.
 OS Arabidopsis thaliana (Mouse-ear cress).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 CC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 CC NCBL_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-cv. Columbia;
 RX MEDLINE-95232183; PubMed-7716229;
 RA Weaver L.M., Lebrun L., Franklin A., Huang L., Hoffman N.,
 RA Wurtele E.S., Nikolau B.J.;
 RT "Molecular cloning of the biotinylated subunit of 3-methylcrotonyl-
 coenzyme A carboxylase of Arabidopsis thaliana.";
 RL Plant Physiol. 107:1013-1014(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-cv. Columbia;
 RX MEDLINE-21016719; PubMed-11130712;
 RA White O., Alonso J., Altarfi H., Araujo R., Bowman C.L., Brooks S.Y.,
 RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
 RA Chung M.K., Conn L., Conway A.B., Cressy T.H., Dewar K.,
 RA Dunn P., Elgu P., Feldblum T.V., Feng J.-D., Fong B., Fujil C.Y.,
 RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Hultzar L.,
 RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
 RA Kim C.J., Koo H.L., Kremetska I., Kuritz D.B., Kwan A., Lam B.,
 RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
 RA Lin X., Liu S.X., Liu Z.A., Lueros J.S., Maltz R., Marziani A.,
 RA Miltscher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.T.,
 RA Pal G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
 RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
 RA Sun H., Taitell L.J., Tambunga G., Tortum M.J., Town C.D.,
 RA Utterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
 RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
 RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
 thaliana.";
 RL Nature 408:816-820(2000).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-cv. Columbia;
 RA Shinozaki K., Davis R.W., Ecker J.R., Theologis A.;
 RT "RIKEN Arabidopsis full length cDNA clones (RAPFs) sequenced by the
 RT SSP consortium (Salk/Stanford/PGEC).";
 RL Submitted (DEC-2001) to the EMBL/Genbank/DBJ databases.
 RN [4]
 RP TISSUE SPECIFICITY.
 RC STRAIN-cv. Landsberg erecta, and cv. Columbia;
 RX MEDLINE-20148760; PubMed-10681539;
 RA McKean A.L., Ke J., Song J., Che P., Achenbach S., Nikolau B.J.,
 RA Wurtele E.S.;
 RT "Molecular characterization of the non-biotin-containing subunit of
 RT 3-methylcrotonyl-CoA carboxylase.";
 RL J. Biol. Chem. 275:5582-5590(2000).
 CC -1- CATALYTIC ACTIVITY: ATP + 3-methylcrotonyl-CoA + HCO(3)(-) - ADP +
 CC phosphate + 3-methylglutaconyl-CoA.
 CC -1- COFACTOR: Biotin.
 CC -1- PATHWAY: leucine catabolism.
 CC -1- SUBUNIT: Probably a heterodimer composed of biotin-containing
 CC alpha subunits and beta subunits (By similarity).
 CC -1- SUBCELLULAR LOCATION: Mitochondrial matrix.
 CC -1- TISSUE SPECIFICITY: In roots, cotyledons, leaves, flowers,
 CC ovaries, siliques and embryos.
 CC MISCELLANEOUS: Temporal and spatial accumulation of the alpha and
 CC beta subunits during development at approximately equal molar
 CC ratios.
 CC -1- CAUTION: Ref. 2 sequence differs from that shown due to erroneous

gene model prediction.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL; U12536; AAA67356.1; -
 DR EMBL; AC006550; AAD25800.1; ALT_SEQ.
 DR EMBL; AY070723; AAL50065.1; -
 DR HSSP; P24182; IBNC.
 DR InterPro; IPR001882; Biotin_attach.
 DR InterPro; IPR000089; Biotin_lipoyl.
 DR InterPro; IPR000901; CPSase.
 DR Pfam; PF02785; Biotin_carb_C; 1.
 DR Pfam; PF00364; Biotin_lipoyl; 1.
 DR Pfam; PF00289; CPSase_L_chain; 2.
 DR Pfam; PF02786; CPSase_L_D2; 2.
 DR PROSITE; PS00188; BIOTIN; 1.
 DR PROSITE; PS00866; CPSASE_1; 1.
 DR PROSITE; PS00867; CPSASE_2; 1.
 DR Mitochondrion; Ligase; Biotin; ATP-binding; Transit peptide.
 DR TRANSIT 1 25 MITOCHONDRION (POTENTIAL)
 FT CHAIN 26 734 METHYLCHROMONYL COA CARBOXYLASE ALPHA
 FT NP_BIND 198 203 ATP (POTENTIAL).
 FT ACT_SITE 329 329 BY SIMILARITY.
 FT BINDING 699 699 BIOTIN (BY SIMILARITY).
 FT CONFLICT 85 85 V -> D (IN REF. 1).
 FT CONFLICT 92 92 A -> AK (IN REF. 1).
 FT CONFLICT 281 300 MISSING (IN REF. 1).
 FT CONFLICT 430 430 W -> L (IN REF. 1).
 SQ SEQUENCE 734 AA; 80451 MW; 251CACE6464B046B CRC64;

Query Match 23.3%; Score 143; DB 1; Length 734;
 Best Local Similarity 45.6%; Pred. No. 0.00056;
 Matches 31; Conservative 10; Mismatches 27; Indels 0; Gaps 0;

QY 54 GEIPAPLAGTGVSKLVKSGDTVRAGQTVLVLEAKMETEINAPDGVKVERVAVGGGGL 113
 DB 665 GTIVAPMAGLVKLVLENKAKVDGQGLIVLEAKMKHEHYKAPSSGSDQLKAKAGQVS 724
 QY 114 GGGLIKI 121
 DB 725 DGSALFRI 732

RESULT 11
 PCRB_METTH STANDARD; PRT; 568 AA.
 ID PCRB_METTH
 AC 027179;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Pyruvate carboxylase subunit B (EC 6.4.1.1) (Pyruvic carboxylase B).
 GN PCRB OR MTH1107.
 OS Methanobacterium thermoautotrophicum.
 OC Archaea; Euryarchaeota; Methanobacteriales;
 CC Methanobacteriaceae; Methanobacter.
 OX NCBI_TaxID=187420;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Delta H;
 RX MEDLINE=98037514; PubMed=9371463;
 RA Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,
 RA Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,
 RA Harrison D., Hoang L., Keagle P., Lum W., Pothier B., Qiu D.,
 RA Spadafora R., Viacre R., Wang Y., Wleczkowski J., Gibson R.,
 RA Uvanti N., Caruso A., Bush D., Sater H., Patwell D., Prabhakar S.,
 RA McDougall S., Shimer G., Goyal A., Pietrovski S., Church G.M.,

RA Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;
 RT "Complete genome sequence of Methanobacterium thermoautotrophicum
 RT deltaH: functional analysis and comparative genomics.";
 RL J. Bacteriol. 179:7135-7155(1997).
 RN [2]
 RP SEQUENCE FROM N.A. AND SEQUENCE OF 1-20.
 RC STRAIN-Delta H;
 RX MEDLINE=98148063; PubMed=9478969;
 RA Mukhopadhyay B., Stoddard S.F., Wolfe R.S.;
 RT "Purification, regulation, and molecular and biochemical
 RT characterization of pyruvate carboxylase from Methanobacterium
 RT thermoautotrophicum strain deltaH.";
 RL J. Biol. Chem. 273:5155-5166(1998).
 CC -1- FUNCTION: PYRUVATE CARBOXYLASE CATALYZES A 2-STEP REACTION,
 CC INVOLVING THE ATP-DEPENDENT CARBOXYLATION OF THE COVALENTLY
 CC ATTACHED BIOTIN IN THE FIRST STEP AND THE TRANSFER OF THE CARBOXYL
 CC GROUP TO PYRUVATE IN THE SECOND. THE MAXIMUM ACTIVITY IS AT PH 8
 CC AND 60 DEGREES CELSIUS.
 CC -1- CATALYTIC ACTIVITY: ATP + pyruvate + HCO(3)(-) = ADP + phosphate +
 CC oxaloacetate.
 CC -1- COFACTOR: ATP, MAGNESIUM (OR MANGANESE OR COBALT), PYRUVATE AND
 CC BICARBONATE.
 CC -1- ENZYME REGULATION: INHIBITED BY ADP AND ALPHA-KETOGLUTARATE.
 CC -1- SUBUNIT: HETEROOCTAMER OF FOUR A AND FOUR B SUBUNITS.
 CC -1- SIMILARITY: WITH OTHER BIOTIN CARBOXYLASES, LIPOLAMIDE TRANSFERASES
 CC AND CARBAMYL PHOSPHATE SYNTHETASES.
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL; AE000881; AAB85596.1; -
 DR EMBL; AF039105; AAC12719.1; -
 DR HSSP; P02905; IBDO.
 DR InterPro; IPR001882; Biotin_attach.
 DR InterPro; IPR000089; Biotin_lipoyl.
 DR InterPro; IPR000891; HMGL-like.
 DR InterPro; IPR003379; PYC_OADA.
 DR Pfam; PF00364; Biotin_lipoyl; 1.
 DR Pfam; PF00682; HMGL-like; 1.
 DR Pfam; PF02436; PYC_OADA; 1.
 DR TIGRPFAMs; TIGR01108; Oada; 1.
 DR PROSITE; PS00188; BIOTIN; FALSE_NEG.
 DR Ligase; Multifunctional enzyme; Gluconeogenesis; Magnesium; Pyruvate;
 KW Biotin; Complete proteome.
 FT BINDING 54 69 PYRUVATE.
 FT BINDING 534 534 BIOTIN.
 SQ SEQUENCE 568 AA; 63955 MW; D328715AB0328DBB CRC64;

Query Match 22.9%; Score 141; DB 1; Length 568;
 Best Local Similarity 29.3%; Pred. No. 0.00062;
 Matches 36; Conservative 19; Mismatches 36; Indels 30; Gaps 2;

QY 5 VYVNGTAIVDVV-----DVKSHENPMCTIIFGGGTGAPAPAGAGAGAGAGEEIPA 58
 DB 469 VEVDGGEFEVKVPTGYMTIEEAPEPVDV-----EGAVNS 504
 QY 59 PLAGTGVSKLVKSGDTVRAGQTVLVLEAKMETEINAPDGVKVERVAVGGGGL 118
 DB 505 TMOGMYVKKLVKSGDQVNAVVAVEAKMKEMENDIOTPHGVVERIYTPAGEKVEGTGDI 564
 QY 119 IKI 121
 DB 565 MVI 567

RESULT 12
 BCCP_PORPU STANDARD; PRT; 157 AA.
 ID BCCP_PORPU


```

FT CONFLICT 132 132 D -> E (IN REF. 1).
FT CONFLICT 238 238 N -> K (IN REF. 1).
FT CONFLICT 268 268 L -> F (IN REF. 1).
FT CONFLICT 546 546 S -> C (IN REF. 1).
FT CONFLICT 642 642 N -> T (IN REF. 1).
FT CONFLICT 771 773 GTA -> STR (IN REF. 1).
FT CONFLICT 831 831 W -> R (IN REF. 1).
FT CONFLICT 839 839 S -> P (IN REF. 1).
FT CONFLICT 1001 1001 Y -> N (IN REF. 1).
FT CONFLICT 1155 1155 K -> R (IN REF. 1).
FT CONFLICT 1178 1178 Q -> P (IN REF. 1).
FT CONFLICT 1180 1180 K -> R (IN REF. 1).
SQ SEQUENCE 1180 AA; 13016 MW; AD60DA3A60F5E001 CAC64;

Query Match
Best Local Similarity 22.4%; Score 138; DB 1; Length 1180;
Matches 25; Conservative 19; Mismatches 21; Indels 0; Gaps 0;

QY 55 EIPAPLAGTVSKILVKGSDVYKAGQVTLVLEAMKMETEINAPDQKVEKVLVNERDAVOC 114
Db 1103 QIGAPMAGVILEVAVKVGKSLVKGESIAVLAMKMEVAVSSPADGVQKVDVFIKDGESVDA 1162
QY 115 GGGLI 119
Db 1163 SDLLV 1167

RESULT 14
BCCP_CHLMU STANDARD; PRT; 163 AA.
ID BCCP_CHLMU
AC 09PKRS;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Biotin carboxyl carrier protein of acetyl-CoA carboxylase (BCCP).
GN ACCB OR TC0399.
OS Chlamydia muridarum.
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=83560;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN-MOPN / N199;
MEDLINE=20150255; PubMed=10684935;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA White O., Hickey E.K., Peterson J., Uetlerback T., Berry K., Bass S.,
RA Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
RA Gwin M., Nelson W., Deboy R., Kolonay J., McClarty G., Salzberg S.L.,
RA Eisen J., Fraser C.M.;
*Genome sequences of Chlamydia trachomatis Mopn and Chlamydia
RT pneumoniae AR39.*
RL Nucleic Acids Res. 28:1397-1406(2000).
CC -1- FUNCTION: THIS PROTEIN IS A COMPONENT OF THE ACETYL COENZYME A
CC CARBOXYLASE COMPLEX; FIRST, BIOTIN CARBOXYLASE CATALYZES THE
CC CARBOXYLATION OF THE CARRIER PROTEIN AND THEN THE TRANSCARBOXYLASE
CC TRANSFERS THE CARBOXYL GROUP TO FORM MALONYL-COA (By similarity).
CC -1- PATHWAY: long-chain fatty acid biosynthesis; first step.
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC EMBL: AE002306; AAF39256.1; -
CC DR HSSP: P02905; 1BDO.
CC DR TIGR: TC0399; -
CC DR InterPro: IPR001249; AccCoA_biotinc.
CC DR InterPro: IPR001882; Biotin_attach.
CC DR InterPro: IPR000089; Biotin_lipoyl.
CC Pfam: PF00364; biotin_lipoyl; 1.

```

```

DR PRINTS; PRO1071; ACOABIOFINCC.
DR TIGRPMAS; TIGR00531; BCCP; 1.
DR PROSITE; PS00188; BIOTIN; 1.
DR Fatty acid biosynthesis; Biotin; Complete proteome.
FT BINDING 127 127 BIOTIN (BY SIMILARITY).
SQ SEQUENCE 163 AA; 18084 MW; 8536D5B85CE1D3B CRC64;

Query Match
Best Local Similarity 21.7%; Score 133.5; DB 1; Length 163;
Matches 42; Conservative 21; Mismatches 57; Indels 21; Gaps 4;

QY 2 KLVYVNGTVADVDVYDKSHENPM--GTLFEGGCGGAPAPA-----ACGAGA 48
Db 20 RIARKRGEFELEERDGPNIQEPFVYDNLRFAGFTQERIPSDONLGNPIVKEVEKKE 79
QY 49 GKAGEGE-IPAPLAGTV-----SKILVKGSDVYKAGQVTLVLEAMKMETEINAPDQK 100
Db 80 DKVEDEDFIVSLVGTFTFGAPSPESPFAFYVPGDIVSEDTVCIVEMKRVANVEYKAGMAGR 139
QY 101 VERVLVNERDAVGGGGLIKI 121
Db 140 VEEVLITNGDPVQFGSKLFRI 160

RESULT 15
PCL1_YEAST STANDARD; PRT; 1178 AA.
ID PCL1_YEAST
AC P1154;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Pyruvate carboxylase 1 (EC 6.4.1.1) (Pyruvate carboxylase 1) (PCB 1).
GN PCL1 OR PYV OR YGL062W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=88298805; PubMed=3042770;
RA Lim F., Morris C.P., Occhiodoro F., Wallace J.C.;
RT "Sequence and domain structure of yeast pyruvate carboxylase.*";
RL J. Biol. Chem. 263:11493-11497(1988).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=5288C;
MEDLINE=97377993; PubMed=9234674;
RA Feuerhahn M., de Montigny J., Potier S., Souciet J.-L.;
RT "The characterization of two new clusters of duplicated genes
RT suggests a 'Lego' organization of the yeast Saccharomyces cerevisiae
RT chromosomes.*";
RL Yeast 13:861-869(1997).
CC [3]
CC SEQUENCE OF 1003-1178 FROM N.A.
CC MEDLINE=87241529; PubMed=3036126;
CC Morris C.P., Lim F., Wallace J.C.;
CC "Yeast pyruvate carboxylase: gene isolation.*";
CC Biochem. Biophys. Res. Commun. 145:390-396(1987).
CC -1- FUNCTION: PYRUVATE CARBOXYLASE CATALYZES A 2-STEP REACTION,
CC INVOLVING THE ATP-DEPENDENT CARBOXYLATION OF THE COVALENTLY
CC ATTACHED BIOTIN IN THE FIRST STEP AND THE TRANSFER OF THE
CC CARBOXYL GROUP TO PYRUVATE IN THE SECOND.
CC -1- CATALYTIC ACTIVITY: ATP + pyruvate + HCO(3)(-) = ADP + phosphate +
CC oxaloacetate.
CC -1- COFACTOR: BIOTIN AND ZINC.
CC -1- PATHWAY: GLUCONEOGENESIS.
CC -1- SUBUNIT: HOMOTETRAMER.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: WITH OTHER BIOTIN CARBOXYLASES, LIPOMIDE TRANSFERASES
CC AND CARBAMYL PHOSPHATE SYNTHETASES.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

```

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; J03889; AAA34843.1; -;
 DR EMBL; Z72584; CAA96765.1; -;
 DR PIR; A29233; OYBP.
 DR HSSP; P24182; IBNC.
 DR SGD; S0003030; PYC1.
 DR InterPro; IPR001882; Biotin_attach.
 DR InterPro; IPR000089; Biotin_lipoyl.
 DR InterPro; IPR000901; CPSase.
 DR InterPro; IPR000891; HMGL-like.
 DR InterPro; IPR003379; PYC_OADA.
 DR Pfam; PF00289; CPSase_L_chain; 1.
 DR Pfam; PF00364; biotin_lipoyl; 1.
 DR Pfam; PF00682; HMGL-like; 1.
 DR Pfam; PF02436; PYC_OADA; 1.
 DR Pfam; PF02785; Biotin_carp_C; 1.
 DR Pfam; PF02786; CPSase_L_D2; 1.
 DR TIGRFAMS; TIGR01235; pyruv_carbox; 1.
 DR PROSITE; PS00188; BIOTIN; 1.
 DR PROSITE; PS00866; CPSASE_1; 1.
 DR PROSITE; PS00867; CPSASE_2; 1.
 DR Ligase; Multifunctional enzyme; Biotin; Glucconeogenesis; ATP-binding;
 KW Zinc; Multigene family.
 FT NP_BIND 182 187
 FT ACT_SITE 312 312
 FT BINDING 1135 1135
 FT SIMILAR 160 330
 FT SIMILAR 350 470
 FT SIMILAR 1086 1178
 FT CONFLICT 462 462
 FT CONFLICT 493 493
 FT CONFLICT 595 595
 FT CONFLICT 619 619
 FT CONFLICT 664 664
 FT CONFLICT 772 772
 FT CONFLICT 879 879
 FT CONFLICT 909 909
 FT SEQUENCE 1178 AA; 130099 MW; BC7110A8AFB23E04 CRC64;
 Query Match 21.6%; Score 133; DB 1; Length 1178;
 Best Local Similarity 42.2%; Pred. No. 0.0052;
 Matches 27; Conservative 15; Mismatches 22; Indels 0; Gaps 0;

QY 56 IPAPLACTVSKILVKEGDTVAKGTVLVLEAMKMETINAPTDGKVEKVLKENDAVOGG 115
 DB 1103 IGAPMAGVVEVYHENGVEVKKGPVAVLSAMKMEIISPSDGVKVFVSDGENVDSS 1162
 QY 116 OGDI 119
 DB 1163 DLV 1166

RESULT 16
 PYC_PICPA STANDARD; PRT; 1189 AA.
 AC P78992;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Pyruvate carboxylase (EC 6.4.1.1) (pyruvic carboxylase) (PCB).
 GN Pichia pastoris (Yeast).
 OS Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 CC Saccharomycetales; Saccharomycetaceae; Pichia.
 OX NCBI_TaxID=4922;
 RN [1]
 RP SEQUENCE FROM N.A.

RX MEDLINE-98301182; PubMed-9639311;
 RA Menendez J., Delgado J., Gancedo C.;
 RT "Isolation of the Pichia pastoris PYC1 gene encoding pyruvate
 RT carboxylase and identification of a suppressor of the pyc
 RT phenotype";
 RL Yeast 14:647-654(1998).
 CC -1- FUNCTION: PYRUVATE CARBOXYLASE CATALYZES A 2-STEP REACTION,
 CC INVOLVING THE ATP-DEPENDENT CARBOXYLATION OF THE COVALENTLY
 CC ATTACHED BIOTIN IN THE FIRST STEP AND THE TRANSFER OF THE
 CC CARBOXYL GROUP TO PYRUVATE IN THE SECOND (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: ATP + pyruvate + HCO(3)(-) = ADP + phosphate +
 CC oxaloacetate.
 CC -1- COFACTOR: BIOTIN AND ZINC.
 CC -1- PATHWAY: GLUCONEOGENESIS.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- SIMILARITY: WITH OTHER BIOTIN CARBOXYLASES, LIPOMIDE TRANSFERASES
 CC AND CARBAMYL PHOSPHATE SYNTHETASES.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; Y11106; CAA71993.1; -;
 DR HSSP; P24182; IDVL
 DR InterPro; IPR001882; Biotin_attach.
 DR InterPro; IPR000089; Biotin_lipoyl.
 DR InterPro; IPR000901; CPSase.
 DR InterPro; IPR000891; HMGL-like.
 DR InterPro; IPR003379; PYC_OADA.
 DR Pfam; PF00289; CPSase_L_chain; 1.
 DR Pfam; PF00364; biotin_lipoyl; 1.
 DR Pfam; PF00682; HMGL-like; 1.
 DR Pfam; PF02436; PYC_OADA; 1.
 DR Pfam; PF02785; Biotin_carp_C; 1.
 DR Pfam; PF02786; CPSase_L_D2; 1.
 DR TIGRFAMS; TIGR01235; pyruv_carbox; 1.
 DR PROSITE; PS00188; BIOTIN; 1.
 DR PROSITE; PS00866; CPSASE_1; 1.
 DR PROSITE; PS00867; CPSASE_2; 1.
 DR Ligase; Multifunctional enzyme; Biotin; Glucconeogenesis; ATP-binding;
 KW Zinc.
 FT NP_BIND 185 190
 FT ACT_SITE 315 315
 FT BINDING 1140 1140
 FT SEQUENCE 1189 AA; 131400 MW; 8B6E858079657914 CRC64;
 Query Match 21.5%; Score 132; DB 1; Length 1189;
 Best Local Similarity 46.6%; Pred. No. 0.0062;
 Matches 27; Conservative 11; Mismatches 20; Indels 0; Gaps 0;

QY 55 EIPAPLACTVSKILVKEGDTVAKGTVLVLEAMKMETINAPTDGKVEKVLKENDAV 112
 DB 1107 EIGAPMAGVVEVYHENGVEVKKGPVAVLSAMKMEIISPSVAGRICQIVAKENDSV 1164

RESULT 17
 BCCP_CHLTR STANDARD; PRT; 164 AA.
 AC O84125;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Biotin carboxyl carrier protein of acetyl-CoA carboxylase (BCCP).
 GN ACB or C1123.
 OS Chlamydia trachomatis.
 CC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
 OX NCBI_TaxID=813;
 RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN-D/WH-3/CX;
RA MEDLINE-99000809; PubMed-9784136;
RA Stephens R.S., Kaiman S., Lammell C.J., Fan J., Marathe R., Aravind L.,
RA Mitchell W.P., Olinger L., Tatusov R.L., Zhao Q., Koonin E.V.,
RA Davis R.W.;
RT "Genome sequence of an obligate intracellular pathogen of humans:
RT Chlamydia trachomatis";
RL Science 282:754-759(1998).
CC -1- FUNCTION: THIS PROTEIN IS A COMPONENT OF THE ACETYL COENZYME A
CC CARBOXYLASE COMPLEX; FIRST, BIOTIN CARBOXYLASE CATALYZES THE
CC CARBOXYLATION OF THE CARRIER PROTEIN AND THEN THE TRANSCARBOXYLASE
CC TRANSFERS THE CARBOXYL GROUP TO FORM MALONYL-CoA (By similarity).
CC -1- PATHWAY: Long-chain fatty acid biosynthesis; first step.
CC -1- SUBUNIT: HOMODIMER (By similarity).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE001286; AAC67714.1; -
DR HSSP: P02905; 3BDO.
DR PHCT-2DPAGE: 084125; -
DR Interpro: IPR001249; ACQA_biotinCC.
DR Interpro: IPR001882; Biotin_attach.
DR Interpro: IPR000089; Biotin_lipoyl.
DR Pfam: PF00364; biotin_lipoyl.1.
DR PRINTS: PR01071; ACOABIORINC.
DR TIGRfam: TIGR00531; BCCP.1.
DR PROSITE: PS00188; BIOTIN.1.
KW Fatty acid biosynthesis; Biotin; Complete proteome.
FT BINDING 128 128 BIOTIN (By similarity).
SQ SEQUENCE 164 AA; 18198 MW; BD0BAACEBC2D384C CRC64;

Query Match 21.3%; Score 131; DB 1; Length 164;
Best Local Similarity 28.2%; Pred. No. 0.0011;
Matches 40; Conservative 22; Mismatches 58; Indels 22; Gaps 4;

QY 2 KLVATNGTAVDADVDKSHENPM--GTLIFGGGTGAPAPAGAG-----AGAGE 53
Db 20 RIYIKREGLELEERTVPSIOEPFYDNLFQSGERIPFDONGNPVIKESIEKKE 79
QY 54 GEIPA-----PLAGTV-----SKILVKGDTFVAKGQTVLVLAMKETEINAPDGS 99
Db 80 SEAPAGDFTVSPVGTGFSPEAPAFIKPDFTVSEDTVCIVAMKVMNEVKAGMSG 139
QY 100 KYEKVLVKERDAVGGGGLIKI 121
Db 140 RVEELITMGDPVFGSKLFRI 161

RESULT 18
ODP2_ECOLI
ID ODP2_ECOLI STANDARD; PRT; 629 AA.
AC P06959;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Dihydrolipoamide acetyltransferase component of pyruvate dehydrogenase
DE complex (EC 2.3.1.12) (E2).
GN ACET OR B0115.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-K12;
RX MEDLINE-83324434; PubMed-6345153;
RA Stephens P.E., Darlison M.G., Lewis H.M., Guest J.R.;

RT "The pyruvate dehydrogenase complex of Escherichia coli K12.
RT Nucleotide sequence encoding the dihydrolipoamide acetyltransferase
RT component";
RL Eur. J. Biochem. 133:481-489(1983).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-K12 / MG1655;
RX MEDLINE-97426617; PubMed-9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shoa Y.;
RT "The complete genome sequence of Escherichia coli K-12";
RL Science 277:1453-1474(1997).
RN [3]
RP SEQUENCE OF 1-12;
RC STRAIN-K12 / EMG2;
RX MEDLINE-97443975; PubMed-9298646;
RA Link A.J., Robinson K., Church G.M.;
RT Comparing the predicted and observed properties of proteins encoded
RT in the genome of Escherichia coli K-12";
RL Electrophoresis 18:1259-1313(1997).
RN [4]
RP SEQUENCE OF 34-46;
RX MEDLINE-84256520; PubMed-6821375;
RA Hale G., Perham R.N.;
RT "Amino acid sequence around lipoyl acid residues in the pyruvate
RT dehydrogenase multienzyme complex of Escherichia coli";
RL Biochem. J. 187:905-908(1980).
RN [5]
RP MUTAGENESIS OF HIS-602;
RX MEDLINE-90351365; PubMed-2201286;
RA Russel G.C., Guest J.R.;
RT "Overexpression of restructured pyruvate dehydrogenase complexes and
RT site-directed mutagenesis of a potential active-site histidine
RT residue";
RL Biochem. J. 269:443-450(1990).
RN [6]
RP LIPOYLATED DOMAINS STUDIES.
RX MEDLINE-91024917; PubMed-2121129;
RA Ali S.T., Guest J.R.;
RT "Isolation and characterization of lipoylated and unlipoated
RT domains of the E2 subunit of the pyruvate dehydrogenase complex of
RT Escherichia coli";
RL Biochem. J. 271:139-145(1990).
CC -1- FUNCTION: The pyruvate dehydrogenase complex catalyzes the overall
CC conversion of pyruvate to acetyl-CoA and CO(2). It contains
CC multiple copies of three enzymatic components: pyruvate
CC dehydrogenase (E1), dihydrolipoamide acetyltransferase (E2) and
CC lipoylated dehydrogenase (E3).
CC -1- CATALYTIC ACTIVITY: Acetyl-CoA + dihydrolipoamide -> CoA + S-
CC acetyldihydrolipoamide.
CC -1- COFACTOR: THE E2 COMPONENT CONTAINS THREE COVALENTLY-BOUND LIPOYL
CC COFACTORS.
CC -1- SUBUNIT: FORMS A 24-POLYPEPTIDE STRUCTURAL CORE WITH OCTAHEDRAL
CC SYMMETRY.
CC -1- SIMILARITY: BELONGS TO THE 2-OXOACID DEHYDROGENASE FAMILY.
CC -1- SIMILARITY: CONTAINS 3 LIPOYL-BINDING DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: V01498; CAA24741.1; -
DR EMBL: D26562; BAA05573.1; -
DR EMBL: AE000120; AAC73226.1; -
DR PIR: A30278; XXEDP.
DR PIR: A16026; A16026.
DR PIR: S45194; S45194.

DR HSP: P10802; 1DPC.
 DR SWISS-2DPAGE: P06959; COLI.
 DR ECODBASE: C062.7; 6TH EDITION.
 DR ECODBASE: C070.0; 6TH EDITION.
 DR EcGene: EG10025; acef.
 DR InterPro: IPR001078; 2oxoacid_dh.
 DR InterPro: IPR000089; biotin_lipoyl.
 DR InterPro: IPR004167; E3_binding.
 DR InterPro: IPR003016; Lipoyl.
 DR Pfam: PF00198; 2-oxoacid_dh; 1.
 DR Pfam: PF00364; biotin_lipoyl; 3.
 DR Pfam: PF02817; e3_binding; 1.
 DR ProDom: PD001115; 2oxoacid_dh; 1.
 DR PROSITE: PS00189; Lipoyl; 3.
 DR GlycoLysis; Transferase; Acyltransferase; Repeat; Lipoyl;
 KW Complete proteome.
 FT INT MET 0 0
 FT DOMAIN 1 314 LIPOYL BINDING, ACIDIC.
 FT DOMAIN 316 629 SUBUNIT BINDING, CATALYTIC.
 FT DOMAIN 372 388 HYDROPHOBIC.
 FT DOMAIN 541 566 HYDROPHOBIC.
 FT BINDING 40 40 LIPOYL.
 FT BINDING 143 143 LIPOYL.
 FT BINDING 244 244 LIPOYL.
 FT ACT_SITE 546 546 POTENTIAL.
 FT ACT_SITE 602 602 POTENTIAL.
 FT ACT_SITE 606 606 POTENTIAL.
 FT REPEAT 1 102
 FT REPEAT 103 203
 FT REPEAT 204 313
 FT MTAGEN 602 602
 SQ SEQUENCE 629 AA; 65964 MW; 058751268B2C0C64; H->C: ABOIISHES CATALYTIC ACTIVITY.

Query Match 21.1%; Score 130; DB 1; Length 629;
 Best Local Similarity 31.6%; Pred. No. 0.0048;
 Matches 43; Conservative 20; Mismatches 53; Indels 20; Gaps 4;
 QY 1 MKLRTVNGTAYDVVDV-DKSHENPMGTILFGGTTGAPAPAGAGAGAGEIPAP 59
 DB 146 MEVAPAPAGATVKEIKVNGDKVSTGLIMFEVAGEAQAAPAKQEAAPAAA----PAP 201
 QY 60 LAGT-----VSKILKEDGTVAAGTIVLLEAMKMETEINAPTDGKVEKVL 105
 DB 202 AAGKEVAVNPDIGDEVEVEVMKVGDKVAEQLTIVEDKASMEVAPFAGVVELK 261
 QY 106 VKERDAVGGGGLIKI 121
 DB 262 VVNGDKVKTG-SLIMI 276

RESULT 19
 MCCA_SOVBN STANDARD; PRT; 731 AA.
 ID MCCA_SOVBN
 AC 0A2777: 042778: Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DE Methylocrotonyl-CoA carboxylase alpha chain, mitochondrial precursor
 DE (RC 6.4.1.4) (3-Methylcrotonyl-CoA carboxylase 1) (MCCase alpha
 DE subunit) (3-methylcrotonyl-CoA:carbon dioxide ligase alpha subunit).
 GN MCCA.
 OS Glycine max (Soybean).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eucots I; Fabales; Fabaceae; Papilionoideae; Phaseolaeae; Glycine.
 OX NCBI_TaxID=3847;
 RN [1]
 RP SEQUENCE FROM N.A.A. AND SEQUENCE OF N-TERMINUS.
 RC STRAIN=cv. Corsoy 79; TISSUE=Coyledon;
 RX MEDLINE=94286521; PubMed=8016064;
 RA Song J., Murele E.S., Nikolau B.J.;
 RT "Molecular cloning and characterization of the cDNA coding for the
 RT biotin-containing subunit of 3-methylcrotonyl-CoA carboxylase;

RT Identification of the biotin carboxylase and biotin-carrier
 RT domains.
 RL Proc. Natl. Acad. Sci. U.S.A. 91:5779-5783(1994).
 CC -1- CATALYTIC ACTIVITY: ATP + 3-methylcrotonyl-CoA + HCO(3)(-) = ADP +
 CC phosphate + 3-methylglutacetyl-CoA.
 CC -1- COFACTOR: Biotin.
 CC -1- PATHWAY: Leucine catabolism.
 CC -1- SUBUNIT: Probably a heterodimer composed of biotin-containing
 CC alpha subunits and beta subunits (By similarity).
 CC -1- SUBCELLULAR LOCATION: Mitochondrial matrix.
 CC -1- TISSUE SPECIFICITY: In leaves, cotyledons and stems.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: U08469; AAA53140.1; -.
 DR EMBL: U08846; AAA53141.1; -.
 DR HSP: P24182; 1BNC.
 DR InterPro: IPR001882; Biotin_attach.
 DR InterPro: IPR000089; Biotin_lipoyl.
 DR InterPro: IPR000901; CPSase.
 DR Pfam: PF02785; Biotin_card_C; 1.
 DR Pfam: PF00364; biotin_lipoyl; 1.
 DR Pfam: PF00289; CPSase_L_chain; 1.
 DR Pfam: PF02786; CPSase_L_D2; 1.
 DR PROSITE: PS00188; BIOTIN; 1.
 DR PROSITE: PS00866; CPSASE_1; 1.
 DR PROSITE: PS00867; CPSASE_2; 1.
 DR Mitochondrion; Ligase; Biotin; ATP-binding; Transit peptide.
 KW Mitochondrion; Ligase; Biotin; ATP-binding; Transit peptide.
 FT TRANSIT 1 22
 FT CHAIN 23 731
 FT NP_BIND 193 198 ATP (POTENTIAL).
 FT ACT_SITE 324 324 BY SIMILARITY.
 FT BINDING 695 695 BIOTIN (BY SIMILARITY).
 FT CONFLICT 68 68 R -> K (IN REF. 1; AAA53141).
 FT CONFLICT 75 75 T -> S (IN REF. 1; AAA53141).
 FT CONFLICT 78 78 E -> K (IN REF. 1; AAA53141).
 SQ SEQUENCE 731 AA; 80619 MW; C4D5A94F8123A964; CRC64;

Query Match 21.1%; Score 130; DB 1; Length 731;
 Best Local Similarity 42.0%; Pred. No. 0.0055;
 Matches 29; Conservative 12; Mismatches 28; Indels 0; Gaps 0;
 QY 53 EGEIPAPLAGTIVSKILVKEGDTVAAGTIVLLEAMKMETEINAPTDGKVEKRDV 112
 DB 660 QGTAVAPAGLVKVLVLENKRVVEGGQVVLLEAMKMEHVAKVSSGVHGLQLMVGQV 719
 QY 113 QGGGGLIKI 121
 DB 720 SDGSVLEFSV 728
 RESULT 20
 BCCP_ANASP STANDARD; PRT; 161 AA.
 ID BCCP_ANASP
 AC 006881;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Biotin carboxyl carrier protein of acetyl-CoA carboxylase (BCCP).
 GN ACCB OR ALL5057.
 OS Anabaena sp. (Strain PCC 7120).
 OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
 OX NCBI_TaxID=103690;
 RN [1]
 RP SEQUENCE FROM N.A.A. AND SEQUENCE OF 1-9.
 RX MEDLINE=93352435; PubMed=8102363;

RA Gornicki P., Scapino L.A., Haselkorn R.;
 RT "Genes for two subunits of acetyl coenzyme A carboxylase of Anabaena
 RT sp. strain PCC 7120: biotin carboxylase and biotin carboxyl carrier
 RT protein.";
 RL J. Bacteriol. 175:5268-5272(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE:21595285; PubMed:11759840;
 RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
 RA Watanabe A., Itiguchi M., Ishikawa A., Kawashima K., Kimura T.,
 RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
 RA Nakazaki N., Shimpo S., Sugimoto M., Takazawa M., Yamada M.,
 RA Yasuda M., Tabata S.;
 RT "Complete genomic sequence of the filamentous nitrogen-fixing
 RT cyanobacterium Anabaena sp. strain PCC 7120.";
 RL DNA Res. 8:205-213(2001).
 CC -1- FUNCTION: THIS PROTEIN IS A COMPONENT OF THE ACETYL COENZYME A
 CC CARBOXYLASE COMPLEX: FIRST, BIOTIN CARBOXYLASE CATALYZES THE
 CC CARBOXYLATION OF THE CARRIER PROTEIN AND THEN THE TRANSCARBOXYLASE
 CC TRANSFERS THE CARBOXYL GROUP TO FORM MALONYL-COA.
 CC -1- PATHWAY: Long-chain fatty acid biosynthesis; first step.
 CC -1- SUBUNIT: HOMODIMER.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: L14863; AAA74628.1; -
 DR EMBL: AF003598; BAB76756.1; -
 DR HSSP: P02905; 1BD0.
 DR InterPro: IPR001249; AccCoA_biotinCC.
 DR InterPro: IPR001882; Biotin_attach.
 DR InterPro: IPR000089; Biotin_lipoyl.
 DR Pfam: PF00364; biotin_lipoyl; 1.
 DR PRINTS: PR01071; AC0AB10TINC.
 DR TIGRFAMs: TIGR00531; BCCP; 1.
 DR PROSITE: PS00188; BIOTIN; 1.
 DR KW Fatty acid biosynthesis; Biotin; Complete proteome.
 FT INIT MET 0
 FT BINDING 145 145 BIOTIN (BY SIMILARITY).
 SQ SEQUENCE 181 AA; 19049 MW; EBEC7B16CC8225F CRC64;
 Query Match 21.1%; Score 129.5; DB 1; Length 181;
 Best Local Similarity 30.7%; Pred. No. 0.0016;
 Matches 31; Conservative 22; Mismatches 35; Indels 13; Gaps 2;
 QY 34 GTGGAPAPAGAGAGAGAGE-----GEIPAPLAGTVSK-----ILVKGSDTVKAGOT 80
 DB 78 GTSRADHVTSSGSGPCKIIDQKLAEVASPMVGTFVRAPAGAEVFEVGDRIHQGT 137
 QY 81 VYLVEAMKMETENAPTDGKVEKYLKERDANVGSGGLIKI 121
 DB 138 VCLIEAMKMLEADYVSGVLEILVNGEPVEYENPLMKI 178
 RESULT 21
 ODP2_PSEAE STANDARD; PRT; 547 AA.
 AC 05638;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE dihydrolipoamide acetyltransferase component of pyruvate dehydrogenase
 DE complex (EC 2.3.1.12) (E2).
 GN ACER OR ACERB OR PA5016.
 OS Pseudomonas aeruginosa.
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 CC Pseudomonas.
 OX NCBI_TaxID=287;

RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-PAO;
 RX MEDLINE:97315227; PubMed:9171401;
 RA Rae J.L., Cutfield J.F., Lamont I.L.;
 RT "Sequences and expression of pyruvate dehydrogenase genes from
 RT Pseudomonas aeruginosa.";
 RL J. Bacteriol. 179:3561-3571(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 1592 / PAO1;
 RX MEDLINE:20437337; PubMed:10984043;
 RA Stover C.K., Pham X.-Q., Erwin A.L., Mizoguchi S.D., Warren P.,
 RA Hickey M.J., Brinkman F.S.L., Huftagle W.O., Kowalik D.J., Lagrou M.,
 RA Garber R.L., Goltry L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
 RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
 RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
 RA Reiter J., Sater M.H., Hancock R.E.W., Lory S., Olson M.V.;
 RT "Complete genome sequence of Pseudomonas aeruginosa PAO1, an
 RT opportunistic pathogen.";
 RL Nature 406:959-964(2000).
 CC -1- FUNCTION: The pyruvate dehydrogenase complex catalyzes the overall
 CC conversion of pyruvate to acetyl-CoA and CO(2). It contains
 CC multiple copies of three enzymatic components: pyruvate
 CC dehydrogenase (E1), dihydrolipoamide acetyltransferase (E2) and
 CC lipoamide dehydrogenase (E3).
 CC -1- CATALYTIC ACTIVITY: Acetyl-CoA + dihydrolipoamide = CoA + S-
 CC acetylhydrolipoamide.
 CC -1- COFACTOR: THE E2 COMPONENT CONTAINS TWO COVALENTLY-BOUND LIPOYL
 CC COFACTORS (BY SIMILARITY).
 CC -1- SUBUNIT: FORMS A 24-POLYPEPTIDE STRUCTURAL CORE WITH OCTAHEDRAL
 CC SYMMETRY.
 CC -1- SIMILARITY: BELONGS TO THE 2-OXOACID DEHYDROGENASE FAMILY.
 CC -1- SIMILARITY: CONTAINS 2 LIPOYL-BINDING DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: U47920; AAC45354.1; -
 DR EMBL: AE004914; AAG08401.1; -
 DR HSSP: P10802; 1EAF.
 DR InterPro: IPR001078; 2oxoacid_dh.
 DR InterPro: IPR000089; Biotin_lipoyl.
 DR InterPro: IPR004167; E3_binding.
 DR InterPro: IPR003016; Lipoyl.
 DR Pfam: PF00198; 2-oxoacid_dh; 1.
 DR Pfam: PF00364; biotin_lipoyl; 2.
 DR Pfam: PF02817; e3_binding; 1.
 DR Prodom: PD001115; 2oxoacid_dh; 1.
 DR PROSITE: PS00189; LIPOYL; 2.
 DR KW Glycolysis; Transferase; Acyltransferase; Repeat; Lipoyl;
 KW Complete proteome.
 FT BINDING 41 41 LIPOYL (BY SIMILARITY).
 FT BINDING 159 159 LIPOYL (BY SIMILARITY).
 FT ACT SITE 520 520 POTENTIAL.
 FT CONFLICT 225 225 A -> V (IN REF. 1).
 FT CONFLICT 295 301 GGAGATG -> AVAPR (IN REF. 1).
 FT CONFLICT 329 329 MQ -> IE (IN REF. 1).
 SQ SEQUENCE 547 AA; 56709 MW; 24E15CC9A590CB4 CRC64;
 Query Match 21.0%; Score 129; DB 1; Length 547;
 Best Local Similarity 37.0%; Pred. No. 0.005;
 Matches 34; Conservative 15; Mismatches 35; Indels 8; Gaps 2;
 QY 38 APAPA---AGGAGAGAGGGEIPAPLAGTVSK-----ILVKGSDTVKAGQVLYLEAKM 89
 DB 101 APAPAPSESKRAPAPAAASVQDIKVPDIGSAGKANYIEVWVRAGDIVEADQSITLESOKA 160

QY 90 ETEINAPTDKYEKYLKERDAVGGGLIKI 121
 DB 161 SMEIPSPAGVSVSIVKGVGTGLIKL 192

RESULT 22
 PYC_MOUSE STANDARD: PRT; 1178 AA.
 ID PYC_MOUSE
 AC Q05920;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Pyruvate carboxylase, mitochondrial precursor (EC 6.4.1.1) (Pyruvate carboxylase) (PCB).
 GN PC OR PCX.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Adipocyte;
 RA MEDLINE=93189578; PubMed=8446588;
 RT Zhang J., Xia W.L., Brew K., Ahmad F.;
 RT "Adipose pyruvate carboxylase: amino acid sequence and domain structure deduced from cDNA sequencing.";
 RL Proc. Natl. Acad. Sci. U.S.A. 90:1766-1770(1993).
 CC -1- FUNCTION: PYRUVATE CARBOXYLASE CATALYZES A 2-STEP REACTION, INVOLVING THE ATP-DEPENDENT CARBOXYLATION OF THE COVALENTLY ATTACHED BIOTIN IN THE FIRST STEP AND THE TRANSFER OF THE CARBOXYL GROUP TO PYRUVATE IN THE SECOND. CATALYZES IN A TISSUE SPECIFIC MANNER, THE INITIAL REACTIONS OF GLUCOSE (LIVER, KIDNEY) AND LIPID (ADIPOSE TISSUE, LIVER, BRAIN) SYNTHESIS FROM PYRUVATE.
 CC -1- CATALYTIC ACTIVITY: ATP + pyruvate + HCO(3)(-) = ADP + phosphate + oxaloacetate.
 CC -1- COFACTOR: BIOTIN AND MANGANESE.
 CC -1- PATHWAY: GLUCONEOGENESIS AND LIPOGENESIS.
 CC -1- SUBUNIT: HOMOTETRAMER.
 CC -1- SUBCELLULAR LOCATION: Mitochondrial matrix.
 CC -1- TISSUE SPECIFICITY: LIVER, KIDNEY, ADIPOSE TISSUE, LIVER AND BRAIN.
 CC -1- SIMILARITY: WITH OTHER BIOTIN CARBOXYLASES, LIPOAMIDE TRANSFERASES AND CARBAMYL PHOSPHATE SYNTHETASES.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
 CC EMBL; L09192; AAA9737.1;
 DR PIR; A47255; A47255.
 DR HSSP; P24182; 1BNC.
 DR SWISS-2DPAGE; Q05920; MOUSE.
 DR MGD; MGI:97520; PCX.
 DR InterPro; IPR001882; Biotin_attach.
 DR InterPro; IPR000089; Biotin_lipoyl.
 DR InterPro; IPR000091; CPSase.
 DR InterPro; IPR000891; HMGL-like.
 DR InterPro; IPR003379; PYC_OADA.
 DR Pfam; PF00289; CPSase_L_chain; 1.
 DR Pfam; PF00364; Biotin_L_chain; 1.
 DR Pfam; PF00682; HMGL-like; 1.
 DR Pfam; PF02436; PYC_OADA; 1.
 DR Pfam; PF02785; Biotin_carb_C; 1.
 DR Pfam; PF02786; CPSase_L_D2; 1.
 DR TIGRFAMS; TIGR01235; pyruv_carbox; 1.
 DR PROSITE; PS00188; BIOTIN; 1.
 KW Lipase; Multifunctional enzyme; Biotin; Manganese; Gluconeogenesis; ATP-binding; Mitochondrion; Lipid synthesis; Transil peptide.
 FT TRANSIT 1 20 MITOCHONDRION (POTENTIAL).

FT CHAIN 21 1178 PYRUVATE CARBOXYLASE
 FT DOMAIN 21 549 BIOTIN CARBOXYLASE (BY SIMILARITY).
 FT DOMAIN 550 1000 CARBOXYLTRANSFERASE (BY SIMILARITY).
 FT DOMAIN 1096 1178 BIOTIN CARBOXYL CARRIER PROTEIN (BY SIMILARITY).
 FT NP_BIND 198 203 ATP (BY SIMILARITY).
 FT ACT_SITE 328 328 BY SIMILARITY.
 FT BINDING 1144 1144 BIOTIN (BY SIMILARITY).
 SQ SEQUENCE 1178 AA; 129684 MW; 14CEA0F9DA8B127 CRC64;

Query Match 21.0%; Score 129; DB 1; Length 1178;
 Best Local Similarity 40.6%; Pred. No. 0.01; 27; Indels 0; Gaps 0;
 Matches 28; Conservative 14; Mismatches 14;

QY 53 EGEIPAPLAGTVSKIVLEGDTVKAGQTVLYLEANKMETEINAPTDKYEKYLKERDAV 112
 DB 1109 KQIGAPMPGKVIDIKVAAGDKVANGPGLCVLSANKMEIVYSPNEGIRKHYVTKMDTL 1168
 QY 113 QGGGGLIKI 121
 DB 1169 EGDLDLLEI 1177

RESULT 23
 ODP2_ACHLA STANDARD: PRT; 544 AA.
 ID ODP2_ACHLA
 AC P35489;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Dihydrolipoamide acetyltransferase component of pyruvate dehydrogenase complex (EC 2.3.1.12) (E2).
 GN PDHC.
 OS Acholeplasma laidlawii.
 OC Bacteria; Firmicutes; Mollicutes; Acholeplasmatales;
 OC Acholeplasmataceae; Acholeplasma.
 OX NCBI_TaxID=2148;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC MEDLINE=92138635; PubMed=1735725;
 RA Walbrandt P., Tegman V., Jonsson B.-H., Wieslander A.;
 RT "Identification and analysis of the genes coding for the putative pyruvate dehydrogenase enzyme complex in Acholeplasma laidlawii.";
 RL J. Bacteriol. 174:1388-1396(1992).
 CC -1- FUNCTION: The pyruvate dehydrogenase complex catalyzes the overall conversion of pyruvate to acetyl-CoA and CO(2). It contains multiple copies of three enzymatic components: pyruvate dehydrogenase (E1), dihydrolipoamide acetyltransferase (E2) and lipoamide dehydrogenase (E3).
 CC -1- CATALYTIC ACTIVITY: Acetyl-CoA + dihydrolipoamide = CoA + S-acetyldihydrolipoamide.
 CC -1- COFACTOR: CONTAINS TWO COVALENTLY-BOUND LIPOYL COFACTOR (POTENTIAL).
 CC -1- SUBUNIT: FORMS A 24-POLYPEPTIDE STRUCTURAL CORE WITH OCTAHEDRAL SYMMETRY (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE 2-OXOACID DEHYDROGENASE FAMILY.
 CC -1- SIMILARITY: CONTAINS 2 LIPOYL-BINDING DOMAINS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
 CC EMBL; M81753; AA21909.1;
 DR PIR; C42653; C42653.
 DR HSSP; P07016; 1C4T.
 DR InterPro; IPR001078; 2Oxoacid_dh.
 DR InterPro; IPR000089; Biotin_lipoyl.
 DR InterPro; IPR004167; E3_binding.
 DR InterPro; IPR003016; Lipoyl.

DR Pfam: PF00198; 2-oxoacid_dh; 1.
 DR Pfam: PF00364; biotin_lipoyl; 2.
 DR Pfam: PF02817; e3_binding; 1.
 DR ProDom: PD001115; 2-oxoacid_dh; 1.
 DR PROSITE: PS00189; LIPOYL; 2.
 KM Glycolysis; transferase; Acyltransferase; Lipoyl.
 FT BINDING 42 42 LIPOYL (BY SIMILARITY).
 FT BINDING 154 154 LIPOYL (BY SIMILARITY).
 FT ACT_SITE 516 516 POTENTIAL.
 SQ SEQUENCE 544 AA; 57261 MW; 81E92D869CFD5424 CRC64;

Query Match 20.8%; Score 128; DB 1; Length 544;
 Best Local Similarity 29.0%; Pred. No. 0.0059;
 Matches 42; Conservative 23; Mismatches 54; Indels 26; Gaps 4;

QY 2 KLTATVNGTAYDVVDVDRKSHENMGTLF-----GGGTGGA-----PAPAAGGA 46
 DB 46 ELSPVDCIT--VSLGKKEGEIHVGOIITYIDGTPTPAAPAPAPVAPAPAPAPAP 103
 QY 47 GAGRAGEGET-----PAPLAGTISKILVREGDTVAKGQTVLVLEAKMETEINAPT 97
 DB 104 VAAPASGDIYDFEFADIGEGHETLQWNEFKVGDYKKEGETLVVETDKVNAELPSV 163
 QY 98 DGKVEKLVKENDAVGSGGLIKIG 122
 DB 164 DGTILKLGKAEGVYHGETVLIG 188

RESULT 24
 BCCA.MYCTU STANDARD; PRT; 654 AA.
 AC P46401;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Acetyl-/propionyl-coenzyme A carboxylase alpha chain [includes: Biotin carboxylase (EC 6.3.4.14); Biotin carboxyl carrier protein (BCCP)].
 GN ACCAL OR BCCA OR RV2501C OR MT2576 OR MTC10747.07C.
 OS Mycobacterium tuberculosis.
 OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae; Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxId=1773;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=50410;
 RX MEDLINE=94222829; PubMed=7909542;
 RA Norman E., de Smet K.A.L., Stoker N.G., Ratledge C., Wheeler P.R., Dale J.W.;
 RT "Lipid synthesis in mycobacteria: characterization of the biotin carboxyl carrier protein genes from Mycobacterium leprae and M. tuberculosis";
 RL J. Bacteriol. 176:2525-2531(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=H37RV;
 RX MEDLINE=98295987; PubMed=9634230;
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D., Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F., Badcock K., Baaham D., Brown D., Chillingworth T., Connor R., Davies R., Devlin K., Fellwell T., Gentles S., Hamlin N., Holtroyd S., Hornsby T., Jagers K., Krogh A., McLean J., Moule S., Murphy L., Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J., Rutter S., Seeger K., Skelton S., Squares S., Squares J., Sulston J.E., Taylor K., Whitehead S., Barrett B.G.;
 RA "Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence.";
 RT Nature 393:537-544(1998).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CDC 1551 / Oshkosh;
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O., Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E., Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,

RA Delcher A., Utterback T., Weidman J., Khouri H., Gail J., Mikula A., Bisht W.;
 RT "Whole genome comparison of Mycobacterium tuberculosis clinical and laboratory strains";
 RT Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
 RU FUNCTION: THIS PROTEIN CARRIES TWO FUNCTIONS: BIOTIN CARBOXYL CARRIER PROTEIN AND BIOTIN CARBOXYLTRANSFERASE.
 CC CATALYTIC ACTIVITY: ATP + biotin-carboxyl-carrier protein + CO(2)
 CC - ADP + phosphate + carboxybiotin-carboxyl-carrier protein.
 CC - COPFACTOR: BIOTIN.
 CC - PATHWAY: Long-chain fatty acid biosynthesis, first step.
 CC - SUBUNIT: MULTIMER COMPRISED OF 2 DIFFERENT SUBUNITS, THE LARGER ONE (63/64 kDa) HAS BIOTIN CARBOXYLASE AND BIOTIN CARRIER FUNCTIONS, WHILE THE SMALLER SUBUNIT POSSESSES CARBOXYLTRANSFERASE AND SUBSTRATE BINDING ACTIVITY.
 CC - SIMILARITY: NO OTHER BIOTIN-DEPENDENT ENZYMS AND CARBAMOYL-PHOSPHATE SYNTHETASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: Z19549; CAA79609.1; -
 DR EMBL: Z95556; CAB08919.1; -
 DR EMBL: AE007094; AAK46880.1; -
 DR HSSP: P24182; 1BNC.
 DR TIGR: MT2576; -
 DR Tuberculist; RV2501C; -
 DR InterPro: IPR001882; Biotin_attach.
 DR InterPro: IPR000089; Biotin_lipoyl.
 DR InterPro: IPR000901; CPase
 DR Pfam: PR00289; CPase_L_chain; 1.
 DR Pfam: PF00364; biotin_lipoyl; 1.
 DR Pfam: PF02785; Biotin_cabd_C; 1.
 DR Pfam: PF02786; CPase_L_D2; 1.
 DR PROSITE: PS00188; BIOTIN; 1.
 DR PROSITE: PS00866; CPASE_1; 1.
 DR PROSITE: PS00867; CPASE_2; 1.
 DR Fatty acid biosynthesis; Ligase; Biotin; Multifunctional enzyme;
 KW ATP-binding; Complete proteome.
 FT DOMAIN 1 437 BIOTIN CARBOXYLASE.
 FT DOMAIN 588 654 BIOTIN CARBOXYL CARRIER PROTEIN.
 FT NP_BIND 162 167 ATP (BY SIMILARITY).
 FT ACT_SITE 294 294 BY SIMILARITY.
 FT BINDING 620 620 BIOTIN (BY SIMILARITY).
 SQ SEQUENCE 654 AA; 70592 MW; FA0A0A1A46432CABF CRC64;

Query Match 20.4%; Score 125.5; DB 1; Length 654;
 Best Local Similarity 40.8%; Pred. No. 0.011; Mismatches 27; Indels 1; Gaps 1;
 Matches 29; Conservative 14;

QY 51 AGEGETPAPLAGTISKILVREGDTVAKGQTVLVLEAKMETEINAPTDKVEKLVKEND 110
 DB 583 ARAAEVSPMPGSGVIAVVGSSQISAGDVVVVLEAKMKEHSLAEVSGRVQ-VLVSVD 641
 QY 111 AVGGGGLIKI 121
 DB 642 QVKEVGLART 652

RESULT 25
 ODP2.AZOVI STANDARD; PRT; 637 AA.
 AC P10802;
 DT 01-JUL-1989 (Rel. 11, Created)
 DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Dihydropyrimidine acetyltransferase component of pyruvate dehydrogenase complex (EC 2.3.1.12) (82).

OS Azotobacter vinelandii.
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 CC Azotobacter.
 OX NCBI_TaxID=354;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 478;
 RX MEDLINE=88271330; PubMed=3292237;
 RA Hanemaaijer R., Janssen A., de Kok A., Veeger C.;
 RT "The dihydrolipoyltransacetylase component of the pyruvate
 dehydrogenase complex from Azotobacter vinelandii. Molecular cloning
 and sequence analysis.";
 RL Eur. J. Biochem. 174:593-599(1988).
 RN [2]
 RP SEQUENCE OF 1-15 AND 380-415.
 RX MEDLINE=88082750; PubMed=3691494;
 RA Hanemaaijer R., de Kok A., Jolles J., Veeger C.;
 RT "The domain structure of the dihydrolipoyl transacetylase component
 of the pyruvate dehydrogenase complex from Azotobacter vinelandii.";
 RL Eur. J. Biochem. 169:245-252(1987).
 RN [3]
 RP LIPOYL DOMAIN CONFORMATION.
 RX MEDLINE=89052887; PubMed=3191993;
 RA Hanemaaijer R., Vervoort J., Westphal A.H., de Kok A., Veeger C.;
 RT "Mobile sequences in the pyruvate dehydrogenase complex, the E2
 component, the catalytic domain and the 2-oxoglutarate dehydrogenase
 complex of Azotobacter vinelandii, as detected by 600 MHz 1H-NMR
 spectroscopy.";
 RL FEBS Lett. 240:205-210(1988).
 RN [4]
 RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 381-637.
 RX MEDLINE=92196586; PubMed=1549782;
 RA Matvev A., Obmolova G., Schulze E., Kalk R.H., Westphal A.H.,
 de Kok A., Hol W.G.J.;
 RT "Atomic structure of the cubic core of the pyruvate dehydrogenase
 multienzyme complex.";
 RL Science 255:1544-1550(1992).
 RN [5]
 RP STRUCTURE BY NMR OF 1-78.
 RX MEDLINE=94222112; PubMed=8068086;
 RA Berg A., Vervoort J., de Kok A.;
 RT "Sequential 1H and 15N nuclear magnetic resonance assignments and
 secondary structure of the N-terminal lipoyl domain of the
 dihydrolipoyl transacetylase component of the pyruvate dehydrogenase
 complex from Azotobacter vinelandii.";
 RL Eur. J. Biochem. 221:87-100(1994).
 RN [6]
 RP STRUCTURE BY NMR OF 1-78.
 RX MEDLINE=97234563; PubMed=9119000;
 RA Berg A., Vervoort J., de Kok A.;
 RT "Three-dimensional structure in solution of the N-terminal lipoyl
 domain of the pyruvate dehydrogenase complex from Azotobacter
 vinelandii.";
 RL Eur. J. Biochem. 244:352-360(1997).
 CC -1- FUNCTION: The pyruvate dehydrogenase complex catalyzes the overall
 conversion of pyruvate to acetyl-CoA and CO(2). It contains
 multiple copies of three enzymatic components: pyruvate
 dehydrogenase (E1), dihydroliponamide acetyltransferase (E2) and
 liponamide dehydrogenase (E3).
 CC -1- CATALYTIC ACTIVITY: Acetyl-CoA + dihydroliponamide -> CoA + S-
 acetyldihydroliponamide.
 CC -1- COFACTOR: THE E2 COMPONENT CONTAINS THREE COVALENTLY-BOUND LIPOYL
 COFACTORS.
 CC -1- SUBUNIT: FORMS A 24-POLYPEPTIDE STRUCTURAL CORE WITH OCTAHEDRAL
 SYMMETRY.
 CC -1- SIMILARITY: BELONGS TO THE 2-OXOACID DEHYDROGENASE FAMILY.
 CC -1- SIMILARITY: CONTAINS 3 LIPOYL-BINDING DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial

CC entitles requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: X12455; CAA30987.1; ALT_INIT.
 CC PIR: S01017; XXAV.
 DR PDB: 1EAA; 31-OCT-93.
 DR PDB: 1EAB; 31-OCT-93.
 DR PDB: 1EAC; 31-OCT-93.
 DR PDB: 1EAD; 31-OCT-93.
 DR PDB: 1EAE; 31-OCT-93.
 DR PDB: 1EAF; 31-OCT-93.
 DR PDB: 1IYU; 12-MAR-97.
 DR PDB: 1IYV; 12-MAR-97.
 DR PDB: 1DPB; 20-APR-95.
 DR PDB: 1DPC; 20-APR-95.
 DR PDB: 1DPD; 20-APR-95.
 DR InterPro: IPR001078; 2-oxoacid_dh.
 DR InterPro: IPR000089; Biotin_lipoyl.
 DR InterPro: IPR004167; E3_binding.
 DR InterPro: IPR003016; Lipoyl.
 DR Pfam: PF00198; 2-oxoacid_dh; 1.
 DR Pfam: PF00364; biotin_lipoyl; 3.
 DR Pfam: PF02817; e3_binding; 1.
 DR ProDom: PD001115; 2-oxoacid_dh; 1.
 DR ProSite: PS00189; Lipoyl; 3.
 KW Glycolysis; Transferrase; Acyltransferase; Repeat; Lipoyl;
 KW 3D-structure.
 FT INIT_MET 0
 FT DOMAIN 1 326
 FT 327 380
 FT DOMAIN 381 637
 FT BINDING 39 39
 FT BINDING 156 156
 FT BINDING 261 261
 FT REPEAT 1 115
 FT REPEAT 116 220
 FT REPEAT 221 326
 FT ACT_SITE 610 610
 FT HELIX 403 406
 FT HELIX 416 431
 FT STRAND 434 442
 FT 444 452
 FT 453 453
 FT 454 459
 FT 460 461
 FT 466 480
 FT 482 484
 FT 486 488
 FT 490 491
 FT 495 497
 FT 503 505
 FT 507 509
 FT 510 511
 FT 512 514
 FT 517 518
 FT 521 523
 FT 526 541
 FT 542 543
 FT 547 550
 FT 555 559
 FT 561 563
 FT 574 575
 FT 582 582
 FT 586 591
 FT 596 609
 FT 610 612
 FT 615 630
 FT 632 636
 FT 636 636
 SO SEQUENCE 637 AA; 64913 MW; D6063B4A5A385F84 CRC64;
 Query Match 20.1%; Score 123.5; DB 1; Length 637;
 Best Local Similarity 36.2%; Pred. No. 0.015; 33; Indels 9; Gaps 2;
 Matches 34; Conservative 13; Mismatches 33;

CC - SPECIFIC MANNER. THE INITIAL REACTIONS OF GLUCOSE (LIVER, KIDNEY)
 CC AND LIPID (ADIPOSE TISSUE, LIVER, BRAIN) SYNTHESIS FROM PYRUVATE.
 CC -1- CATALYTIC ACTIVITY: ATP + pyruvate + HCO(3)(-) = ADP + phosphate +
 CC oxaloacetate.
 CC -1- COFACTOR: BIOTIN AND MANGANESE.
 CC -1- PATHWAY: GLUCONEOGENESIS AND LIPOGENESIS.
 CC -1- SUBUNIT: HOMOTETRAMER.
 CC -1- SUBCELLULAR LOCATION: Mitochondrial matrix.
 CC -1- DISEASE: DEFICIENCY IN PC CAUSES LACTIC ACIDOSIS, MENTAL
 CC RETARDATION AND DEATH. OCCURS IN THREE FORMS: TYPE A (MILD); TYPE
 CC B (SEVERE NEONATAL) AND A VERY MILD LACTICACIDEMIA.
 CC -1- SIMILARITY: WITH OTHER BIOTIN CARBOXYLASES, LIPONAMIDE TRANSFERASES
 CC AND CARBAMYL PHOSPHATE SYNTHETASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U04641; AAA99537.1; -
 CC EMBL: S72370; AAA31500.1; -
 CC EMBL: U30891; AAA82937.1; -
 CC EMBL: M26122; AAA6423.1; -
 CC EMBL: K02282; AAA60033.1; -
 CC PIR: B27883; B27883.
 CC PIR: S01469; S01469.
 CC HSSP: P24182; 1BNC.
 CC Gene: HGNC:8636; PC.
 CC MIM: 266150; -
 CC InterPro: IPR001882; Biotin_attach.
 CC InterPro: IPR000089; Biotin_lipoyl.
 CC InterPro: IPR000901; CPSase.
 CC InterPro: IPR000891; HMGL-like.
 CC InterPro: IPR003379; PYC-OMDA.
 CC Pfam: PF00289; CPSase_L_chain; 1.
 CC Pfam: PF00364; Biotin_lipoyl; 1.
 CC Pfam: PF00682; HMGL-like; 1.
 CC Pfam: PF02436; PYC-OMDA; 1.
 CC Pfam: PF02785; Biotin_card_C; 1.
 CC Pfam: PF02786; CPSase_L_D2; 1.
 CC Pfam: TIGR01235; pyruv_carbox; 1.
 CC PROSITE: PS00188; BIOTIN; 1.
 CC Lfbase: Multifunctional enzyme: Biotin; Manganese; Gluconeogenesis;
 CC ATP-binding; Mitochondrion; Lipid synthesis; Transil peptide;
 CC Disease mutation.
 CC KW TRANSIT 20
 CC KW CHAIN 21
 CC KW DOMAIN 21
 CC KW DOMAIN 550
 CC KW DOMAIN 1096
 CC NP_BIND 198
 CC AC_SITE 328
 CC BINDING 1144
 CC VARIANT 610
 CC VARIANT 743
 CC VARIANT 743
 CC CONFLICT 225
 CC CONFLICT 352
 CC CONFLICT 385
 CC CONFLICT 486
 CC CONFLICT 638
 CC CONFLICT 729
 CC CONFLICT 774
 CC SEQUENCE 1178 AA; 129633 MW; 381F527553A20095 CRC64;
 CC Query Match 19.7%; Score 121; DB 1; Length 1178;
 CC Best Local Similarity 40.6%; Pred. NO. 0.043; 28; Indels 0; Gaps 0;
 CC Matches 28; Conservative 13; Mismatches 28; Indels 0; Gaps 0;

Oy 53 ECEIPAPLAGTVSKILVEGDTVKAGQVYLLEANKMETEINAPDGVKEVLYERDAV 112
 Db 1109 KGQIGAPMPGKVIDIKVVAAGAKVAGKQPLCVLSANKMETVYSPNEGIVRKVHTKDMTL 1168
 Oy 113 QGGGGLINI 121
 Db 1169 EGGDLILEI 1177
 RESULT 28
 PTC_RAT
 ID PTC_RAT STANDARD; PRT; 1178 AA.
 AC P52873; Q64555;
 DT 01-OCT-1996 (Rel. 34, created)
 DT 01-OCT-1996 (Rel. 34, last sequence update)
 DT 15-JUN-2002 (Rel. 41, last annotation update)
 DE pyruvate carboxylase, mitochondrial precursor (EC 6.4.1.1) (Pyruvic
 DE carboxylase) (PCB).
 GN PC.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RX MEDLINE=96096548; PubMed=8522203;
 RA Lehn D.A., Moran S.M., Macdonald M.J.;
 RT "The sequence of the rat pyruvate carboxylase-encoding cDNA.";
 RL Gene 165:331-332(1995).
 RN (2)
 RP SEQUENCE FROM N.A.
 RC STRAIN=Wistar; TISSUE=Liver;
 RX MEDLINE=96257760; PubMed=8687410;
 RA Jitrapakdee S., Booker G.W., Cassidy A.I., Wallace J.C.;
 RT "Cloning, sequencing and expression of rat liver pyruvate
 RT carboxylase.";
 RL Biochem. J. 316:631-637(1996).
 CC -1- FUNCTION: PYRUVATE CARBOXYLASE CATALYZES A 2-STEP REACTION,
 CC INVOLVING THE ATP-DEPENDENT CARBOXYLATION OF THE COVALENTLY
 CC ATTACHED BIOTIN IN THE FIRST STEP AND THE TRANSFER OF THE
 CC CARBOXYL GROUP TO PYRUVATE IN THE SECOND. CATALYZES IN A TISSUE
 CC SPECIFIC MANNER. THE INITIAL REACTIONS OF GLUCOSE (LIVER, KIDNEY)
 CC AND LIPID (ADIPOSE TISSUE, LIVER, BRAIN) SYNTHESIS FROM PYRUVATE.
 CC -1- CATALYTIC ACTIVITY: ATP + pyruvate + HCO(3)(-) = ADP + phosphate +
 CC oxaloacetate.
 CC -1- COFACTOR: BIOTIN AND MANGANESE (BY SIMILARITY).
 CC -1- PATHWAY: GLUCONEOGENESIS AND LIPOGENESIS.
 CC -1- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Mitochondrial matrix.
 CC -1- SIMILARITY: WITH OTHER BIOTIN CARBOXYLASES, LIPONAMIDE TRANSFERASES
 CC AND CARBAMYL PHOSPHATE SYNTHETASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U32314; AAA6256.1; -
 CC EMBL: U36585; AAC52668.1; -
 CC HSSP: P24182; 1BNC.
 CC InterPro: IPR001882; Biotin_attach.
 CC InterPro: IPR000089; Biotin_lipoyl.
 CC InterPro: IPR000901; CPSase.
 CC InterPro: IPR000891; HMGL-like.
 CC Pfam: PF00289; CPSase_L_chain; 1.
 CC Pfam: PF00364; Biotin_lipoyl; 1.
 CC Pfam: PF00682; HMGL-like; 1.

DR Pfam; PF02436; PYC_OADA; 1.
DR Pfam; PF02785; Biotin_carb_C; 1.
DR Pfam; PF02786; CPSase_L_D2; 1.
DR TIGRfams; TIGR01235; Pyruv_carbox; 1.
DR PROSITE; PS00188; BIOTIN; 1.
KW Ligase; Multifunctional enzyme; Biotin; Manganese; Gluconeogenesis;
KM ATP-binding; Mitochondrion; Lipid synthesis; Translt peptide;
FT TRANSIT 1 20
FT CHAIN 21 1178
FT DOMAIN 21 549
FT DOMAIN 550 1000
FT DOMAIN 1096 1178
FT NP_BIND 198 203
FT ACT_SITE 328 328
FT BINDING 1144 1144
FT CONFLICT 222 222
FT CONFLICT 866 866
FT CONFLICT 977 977
SQ SEQUENCE 1178 AA; 129689 MW; 8E5FA19BC132A8DD CRC64;
Query Match 19.3%; Score 119; DB 1; Length 1178;
Best Local Similarity 37.7%; Pred. No. 0.06; Mismatches 0; Gaps 0;
Matches 26; Conservative 15; Mismatches 28; Indels 0;
OY 53 EGEPLAAGTGVSKILYKEDGYKAGQTVLVLEAMKETEINAPDGRVERVYKERDAY 112
DB 1109 KGOIGAPMPKVIDKVAAGAKVKGQPLCVLSAMKMETVVTSPMEGTIRKVTXKMTL 1168
OY 113 OGGGGGLIKI 121
DB 1169 EGGDLLEI 1177
RESULT 29
PCCA_RAT ID STANDARD; PRT; 704 AA.
AC 01-APR-1990 (Rel. 14, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Propionyl-CoA carboxylase alpha chain, mitochondrial precursor
DE (RC 6.4.1.3) (PCCase alpha subunit) (Propanoyl-CoA:carbon dioxide
DE ligase alpha subunit) (Fragment).
GN PCCA.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=89308706; PubMed=2745462;
RX Browner M.F., Taroni F., Szul E., Rosenberg L.E.;
RT "Sequence analysis, biogenesis, and mitochondrial import of the
RT alpha-subunit of rat liver propionyl-CoA carboxylase.";
RL J. Biol. Chem. 264:12680-12685(1989).
RN [2]
RP REVISIONS.
RA Browner M.F., Taroni F., Szul E., Rosenberg L.E.;
RL Submitted (FEB-1989) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: ATP + propanoyl-CoA + HCO(3)(-) - ADP +
CC phosphate + (S)-methylmalonyl-CoA.
CC -1- COFACTOR: BIOTIN.
CC -1- PATHWAY: KEY ENZYME IN THE CATABOLIC PATHWAY OF ODD-CHAIN
CC FATTY ACIDS, ISOLEUCINE, THREONINE, METHIONINE, AND VALINE.
CC -1- SUBUNIT: PROBABLY A DODECAMER COMPOSED OF SIX BIOTIN-CONTAINING
CC ALPHA SUBUNITS AND SIX BETA SUBUNITS.
CC -1- SUBCELLULAR LOCATION: Mitochondrial matrix.
CC -1- DISEASE: PROPIONIC ACIDEMIA DUE TO RECESSIVELY INHERITED
CC DEFICIENCY OF PCCase ACTIVITY OFTEN CAUSES LIFE-THREATENING
CC KETOSIS AND ACIDOSIS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M22631; AAA88512.1; ALT_SEQ.
DR PIR; A34337; A34337.
DR HSSP; P24182; IDV1.
DR InterPro; IPR001882; Biotin_attach.
DR InterPro; IPR000089; Biotin_lipoyl.
DR InterPro; IPR000901; CPSase.
DR Pfam; PF00289; CPSase_L_chain; 1.
DR Pfam; PF00364; biotin_lipoyl; 1.
DR Pfam; PF02785; Biotin_carb_C; 1.
DR Pfam; PF02786; CPSase_L_D2; 1.
DR PROSITE; PS00188; BIOTIN; 1.
DR PROSITE; PS00866; CPSASE_1; 1.
DR PROSITE; PS00867; CPSASE_2; 1.
DR PROSITE; PS00867; CPSASE_2; 1.
KW Mitochondrion; Ligase; Biotin; ATP-binding; Translt peptide.
FT TRANSIT 1 21
FT CHAIN 22 704
FT NP_BIND 199 204
FT ACT_SITE 329 329
FT BINDING 670 670
SQ SEQUENCE 704 AA; 77711 MW; 36CEC52DFD2A8A CRC64;
Query Match 19.3%; Score 118.5; DB 1; Length 704;
Best Local Similarity 27.8%; Pred. No. 0.04; Mismatches 42; Conservative 21; Mismatches 51; Indels 37; Gaps 4;
Matches 42; Conservative 21; Mismatches 51; Indels 37; Gaps 4;
OY 8 NGTAYDVVDVDDKSH-----ENPMGTIFGG--GTGAPAPAGG----- 45
DB 553 NGTFPENVDSKLTNTSTNLAASPLSVNVDSTORTVGCSDADAGNSIOPLGTVYV 612
OY 46 -----AGAGKAGEEIRP-----APLAGTGVSKILYKEDGYKAGQTVLVLEAMKE 90
DB 613 HILTKLAELNKFMLEKVPKDTSSVLRSPKGVAVVSPGDMVAEGEIVCEIAMKKQ 672
OY 91 TELNAPDQKVEVIVKERNVAGGGGLIKI 121
DB 673 NSMTAGKMGKVLHCKACDPTVGGDLVEL 703
RESULT 30
BFB7_MYCSM ID STANDARD; PRT; 70 AA.
AC 09XCD6;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Biotinylated protein BFB7.3 homolog.
DE Mycobacterium smegmatis.
OS Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1772;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=ATCC 700084 / mc(2)155;
RX MEDLINE=9328972; PubMed=10400584;
RA Fernandes N.D., Wu Q.-L., Kong D., Puyang X., Garg S., Husson R.N.;
RT "A mycobacterial extracytoplasmic sigma factor involved in survival
RT following heat shock and oxidative stress.";
RL J. Bacteriol. 181:4266-4274(1999).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC : or send an email to license@slb-sib.ch).

CC :
 DR EMBL; AF144091; AAD41812.1;
 DR HSSP; P10802; IYU.
 DR InterPro; IPR001882; Biotin_attach.
 DR InterPro; IPR000089; Biotin_1poyl.
 DR Pfam; PF00364; Biotin_1poyl; 1.
 DR PROSITE; PS00188; BIOTIN; FALSE_NEG.
 KW Biotin.
 FT INIT_MET 0 0 BY SIMILARITY.
 FT BINDING 36 36 BIOTIN (BY SIMILARITY).
 SO SEQUENCE 70 AA; 7306 MW; 975C293B63C770C8 CRC64;

Query Match 19.2%; Score 118; DB 1; Length 70;
 Best Local Similarity 39.3%; Pred. No. 0.0047;
 Matches 24; Conservative 14; Mismatches 23; Indels 0; Gaps 0;

QY 55 EIPAPLACTVSKILYKEDDTYKAGQTVLYLEAMKMETEINAFPTDCKVEKVLKERNDAVQG 114
 Db 3 DVRAEIVASVLEVVVHEDQIGEGDTLVLLSEMKMEIPVLAEVAGTVTKVNVVAEGDVIOA 62
 QY 115 G 115
 Db 63 G 63

Search completed: May 1, 2003, 07:52:45
 Job time : 14.8021 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 1, 2003, 07:49:48 ; Search time 10.5729 Seconds

(without alignments)
636.476 Million cell updates/sec

Title: US-09-987-485-2

Perfect score: 342
Sequence: 1 ESEIPAPLAGTVSKILVREG.....KVLVNERDAVGGGGLIKIG 70Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database :

1: PIR_73:*
2: PIR1:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	342	100.0	123	1 BKIP	biotin carboxyl ca
2	195	57.0	145	2 F75135	methyilmalonyl-coa
3	189	55.3	149	2 A71074	probable methyilmal
4	180	52.6	571	2 F71133	probable oxaloacet
5	178	52.0	140	2 H69526	methyilmalonyl-coa
6	174	50.9	655	2 B70432	pyruvate carboxyla
7	174	50.9	984	2 T44608	pyruvate carboxyla
8	173	50.6	567	2 F64453	oxaloacetate decar
9	170	49.7	186	2 D90418	hypothetical prote
10	167	48.8	134	2 C72341	propionyl-CoA carb
11	167	48.8	1144	2 D97227	pyruvate carboxyla
12	165	48.2	536	2 A28088	oxaloacetate decar
13	164	48.0	597	2 G82308	oxaloacetate decar
14	163	47.7	599	2 D81367	probable pyruvate
15	161	47.1	142	2 D69510	oxaloacetate decar
16	159.5	46.6	576	2 H98347	biotin carboxylase
17	159.5	46.6	576	2 AC3038	biotin carboxylase
18	158	46.5	665	2 G97819	hypothetical prote
19	158	46.2	607	2 F82966	probable transcarb
20	157	45.9	1146	2 AC1565	pyruvate carboxyla
21	157	45.9	1146	2 AH1208	pyruvate carboxyla
22	157	45.9	1150	2 A83978	pyruvate carboxyla
23	153	44.7	620	2 F70439	oxaloacetate decar
24	152	44.4	167	2 T44291	biotin carboxyl ca
25	152	44.4	436	2 A53568	methyilmalonyl-Co
26	152	44.4	591	2 B44465	sodium ion pump ox
27	150	43.9	70	2 E83791	hypothetical prote
28	150	43.9	665	2 C71667	propionyl-CoA carb
29	148	43.3	591	2 AB0509	oxaloacetate decar

30	148	43.3	591	2 AE0909	oxaloacetate decar
31	148	43.3	1174	2 C97686	pyruvate carboxyla
32	148	43.3	1174	2 AE2911	pyruvate carboxyla
33	146	42.7	1150	2 G89881	pyruvate carboxyla
34	145	42.4	655	2 AB3395	probable acyl-CoA
35	144	42.1	161	2 A95049	hypothetical prote
36	144	42.1	161	2 G97919	acetyl-CoA carboxy
37	143	41.8	730	2 G86161	hypothetical prote
38	143	41.8	1137	2 E86708	pyruvate carboxyla
39	143	41.8	1185	2 T39734	pyruvate carboxyla
40	142	41.5	155	2 H86721	hypothetical prote
41	141	41.2	1078	2 D87647	hypothetical prote
42	140	40.9	568	2 C69014	oxaloacetate decar
43	139.5	40.8	1127	2 D70671	pyruvate carboxyla
44	138.5	40.5	157	2 S73204	acetyl-CoA carboxy
45	138	40.4	129	2 D49094	methyilmalonyl-coa
46	138	40.4	1180	2 S46094	pyruvate carboxyla
47	136	39.8	134	2 T44984	pyruvate carboxyla
48	135	39.5	661	2 F83284	probable biotin ca
49	133.5	39.0	159	2 D97338	biotin carboxyl ca
50	133	38.9	1178	1 QYBP	pyruvate carboxyla
51	132.5	38.7	162	2 A97521	biotin carboxyl ca
52	132.5	38.7	162	2 AC2740	hypothetical prote
53	132.5	38.7	170	2 F87482	hypothetical prote
54	130.5	38.2	654	2 G87517	acetyl/propionyl-C
55	130	38.0	598	2 G86699	hypothetical prote
56	130	38.0	598	2 A55579	biotin carboxyl ca
57	130	38.0	731	2 T06360	biotin carboxyl ca
58	129	37.7	667	2 C70719	biotin carboxylase
59	129	37.7	1148	2 F69685	pyruvate carboxyla
60	129	37.7	1178	1 A47255	pyruvate carboxyla
61	128.5	37.6	593	2 B71373	probable oxaloacet
62	128.5	37.6	610	2 G84306	biotin carboxylase
63	127	37.1	600	2 F70980	probable accA3 pro
64	126.5	37.0	182	2 B53311	acetyl-CoA carboxy
65	126.5	37.0	182	2 A12437	biotin carboxyl ca
66	126	36.8	678	2 AG3492	propionyl-CoA carb
67	124.5	36.4	156	2 AH3384	biotin carboxyl ca
68	124.5	36.4	654	2 B55579	biotin carboxyl ca
69	124	36.3	665	2 T02793	propionyl-CoA carb
70	122	35.7	164	2 G71553	probable biotin ca
71	121	35.4	1178	1 JC2460	pyruvate carboxyla
72	120.5	35.2	187	2 H87475	urea amidolysase-re
73	120.5	35.2	187	2 C75558	acetyl-CoA carboxy
74	120	35.1	724	2 G89932	hypothetical prote
75	119.5	34.9	154	2 T16187	hypothetical prote
76	119.5	34.9	158	2 T30379	biotin carboxyl ca
77	119	34.8	591	2 S71009	biotin carboxylase
78	119	34.8	597	2 S71006	biotin carboxylase
79	119	34.8	1158	2 AC3285	pyruvate carboxyla
80	119	34.8	80	2 JC4391	pyruvate carboxyla
81	118	34.5	163	2 C81708	acetyl-coenzyme A
82	118	34.5	688	2 T21641	hypothetical prote
83	117	34.2	410	2 H83877	acetoic dehydrogen
84	117	34.2	662	2 G95929	probable methylo
85	117	34.2	670	2 H96022	probable propionyl
86	117	34.2	1095	2 B83471	probable pyruvate
87	116.5	34.1	154	2 D70418	biotin carboxyl ca
88	116	33.9	590	2 T42202	probable acyl-CoA
89	116	33.9	590	2 T42206	probable acyl-CoA
90	116	33.9	663	2 AF2984	3-methylcrotonyl-C
91	116	33.9	709	2 A98299	3-methylcrotonyl-C
92	115.5	33.8	547	2 H83018	3-methylcrotonyl-C
93	115	33.6	71	2 F70596	hypothetical prote
94	115	33.6	73	2 T44814	hypothetical prote
95	114.5	33.5	630	2 G85494	hypothetical prote
96	114.5	33.5	630	2 G90643	hypothetical prote
97	113.5	33.2	630	1 XXECDP	hypothetical prote
98	111.5	32.6	151	2 E81033	acetyl-CoA carboxy
99	111.5	32.6	1195	2 T43735	pyruvate carboxyla
100	111	32.5	704	2 A34337	propionyl-CoA carb

ALIGNMENTS

RESULT 1

BKIP

biofin carboxyl carrier protein [validated] - *Propionibacterium freudenreichii* subsp. sh
C:Species: *Propionibacterium freudenreichii* subsp. *Shermanti*
C:Date: 31-Mar-1980 #sequence_revision 31-Mar-1980 #text_change 01-Feb-2002
C:Accession: A03401
R:Maloy, M. L., Bowler, B. D., Zwolsky, G. K., Kumar, K. G., Wood, H. G., Ericsson, L. H., W
A:Title: Amino acid sequence of the biotinyl subunit from transcarboxylase.
A:Reference number: A03401; M01D:80049796; PMID:40965
A:Accession: A03401
A:Molecule type: protein
A:Residues: 1-123 <MAL>
C:Comment: Six or 12 chains of biotin carboxyl carrier protein (BCCP) are found in the t
onyl coenzyme A to BCCP and (2) from BCCP to pyruvate, forming oxalacetate.
C:Comment: See PIR:A4865 and PIR:S36808.
C:Keywords: biotin carboxyl carrier protein; lipoyl/biotin-binding homology
F:50-123/Domain: lipoyl/biotin-binding homology <LPS>
F:89/Binding site: biotin (Lys) (covalent) #status experimental

Query Match 100.0%; Score 342; DB 1; Length 123;
Best Local Similarity 100.0%; Pred. No. 2.1e-27;
Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EGEIAPLAGVSKILVEGDTVKAQGVTLVLEAMKMETEINAPTDGVEKLVKERDAV 60

DB 54 EGEIAPLAGVSKILVEGDTVKAQGVTLVLEAMKMETEINAPTDGVEKLVKERDAV 113

OY 61 OGCGGLIKIG 70
|||||||
DB 114 OGCGGLIKIG 123

RESULT 2

methyilmalonyl-CoA decarboxylase gamma chain PAB1771 - *Pyrococcus abyssi* (strain Orsay)
C:Species: *Pyrococcus abyssi*
C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Jun-2000
C:Accession: F75135
R:anonymous; Genoscope
A:Description: *Pyrococcus abyssi* genome sequence: Insights into archaeal chromosome stru
A:Reference number: A75001
A:Accession: F75135
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-145 <KAW>
A:Cross-references: GB:AJ248285; GB:AL096836; NID:g5458067; PIDN:CA849799.1; PID:g545831
A:Experimental source: strain Orsay
C:Genetics:
A:Gene: PAB1771
C:Superfamily: biotin carboxyl carrier protein; lipoyl/biotin-binding homology

Query Match 57.0%; Score 195; DB 2; Length 145;
Best Local Similarity 58.6%; Pred. No. 1.4e-12;
Matches 41; Conservative 10; Mismatches 19; Indels 0; Gaps 0;

OY 1 EGEIAPLAGVSKILVEGDTVKAQGVTLVLEAMKMETEINAPTDGVEKLVKERDAV 60

DB 76 EGVVAPAPMGKVLKILVGGQGVKIGGILLLEAMKMEINAPRPGVVKRILVKEGDAV 135

OY 61 OGCGGLIKIG 70
|||||
DB 136 DTGAPLIEIG 145

RESULT 3

A71074

probable methyilmalonyl-CoA decarboxylase gamma chain - *Pyrococcus horikoshii*

C:Species: *Pyrococcus horikoshii*
C:Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 20-Jun-2000
C:Accession: A71074
R:Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Halkawa, Y.; Hino, Y.; Yamamoto, S.; Se
M.; Ohnuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Ogu
DNA Res. 5, 55-76, 1998
A:Title: Complete sequence and gene organization of the genome of a hyper-thermophilic
A:Reference number: A71000; M01D:98344137; PMID:9679194
A:Accession: A71074
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-149 <KAW>
A:Cross-references: GB:AP000005; NID:g3236132; PIDN:BAA30387.1; PID:g3237704
A:Experimental source: strain OT3
A:Note: this accession replaces an interim accession for a sequence replaced by Genba
C:Genetics:
A:Gene: PH1284
C:Superfamily: biotin carboxyl carrier protein; lipoyl/biotin-binding homology
F:76-149/Domain: lipoyl/biotin-binding homology <LPS>

Query Match 55.3%; Score 189; DB 2; Length 149;
Best Local Similarity 54.3%; Pred. No. 5.8e-12;
Matches 38; Conservative 14; Mismatches 18; Indels 0; Gaps 0;

OY 1 EGEIAPLAGVSKILVEGDTVKAQGVTLVLEAMKMETEINAPTDGVEKLVKERDAV 60

DB 80 EGVVAPAPMGKVLKILVGGQGVKIGGILLLEAMKMEINAPRPGVVKRILVKEGDAV 139

OY 61 OGCGGLIKIG 70
|||||
DB 140 DTGAPLIEIG 149

RESULT 4

probable oxalacetate decarboxylase alpha chain - *Pyrococcus horikoshii*
C:Species: *Pyrococcus horikoshii*
C:Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 18-Aug-2000
C:Accession: F71133
R:Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Halkawa, Y.; Hino, Y.; Yamamoto, S.; Se
M.; Ohnuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Ogu
DNA Res. 5, 55-76, 1998
A:Title: Complete sequence and gene organization of the genome of a hyper-thermophilic
A:Reference number: A71000; M01D:98344137; PMID:9679194
A:Accession: F71133
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-571 <KAW>
A:Cross-references: GB:AP000003; NID:g3236130; PIDN:BAA29928.1; PID:g3237245
A:Experimental source: strain OT3
A:Note: this accession replaces an interim accession for a sequence replaced by Genba
C:Genetics:
A:Gene: PH0834
C:Superfamily: *Klebsiella pneumoniae* oxalacetate decarboxylase alpha chain; lipoyl/b
F:496-571/Domain: lipoyl/biotin-binding homology <LPS>
F:537/Binding site: biotin (Lys) (covalent) #status predicted

Query Match 52.6%; Score 180; DB 2; Length 571;
Best Local Similarity 54.5%; Pred. No. 1.9e-10;
Matches 36; Conservative 14; Mismatches 16; Indels 0; Gaps 0;

OY 4 IPAPLAGVSKILVEGDTVKAQGVTLVLEAMKMETEINAPTDGVEKLVKERDAV 63

DB 505 VSAPAPMGKVLKILVGGQGVKIGGILLLEAMKMEINAPRPGVVKRILVKEGDAV 564

OY 64 OGGLIKI 69
|||||
DB 565 QPLIEL 570

RESULT 5

H69526

A:Reference number: A96900; MUID:21359325; PMID:21359325
A:Accession: D97227
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1144 <NR>
A:Cross-references: GB:AE001437; PIDN:AAK80607.1; PTD:g15025689; GSPDB:GN00168
A:Experimental source: Clostridium acetobutylicum ATCC824
C:Genetics:
A:Gene: CAC2660
C:Superfamily: pyruvate carboxylase; biotin carboxylase homology; lipoYL/biotin-binding

Query Match 48.8%; Score 167; DB 2; Length 1144;
Best Local Similarity 50.7%; Pred. No. 7, Be-09;
Matches 34; Conservative 11; Mismatches 22; Indels 0; Gaps 0;

OY 3 EIPPLACTGVSKILVKEGDFVAKQTVLVLEAMKETETINAPTDRKVEKVLVKERDAVQG 62
|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 1077 EIGASIPENVNVKVFVKPGDKRVKGDSLWIVIAMKMETNVSSEDTGVGIVFKEGDVGVS 1136

OY 63 GGGIIRKI 69
|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 1137 GOLLVLT 1143

RESULT 12
A28088
oxaloacetate decarboxylase (EC 4.1.1.3) alpha chain - Klebsiella pneumoniae
N:Alternate names: oxalate beta-decarboxylase, alpha chain
C:Species: Klebsiella pneumoniae
C>Date: 19-May-1989 #sequence_revision 19-May-1989 #text_change 11-Jan-2002
C:Accession: A28088
R:Schwartz, E.; Oesterhelte, D.; Reinke, H.; Beyreuther, K.; Dimroth, P.
J. Biol. Chem. 263, 9640-9645, 1988
A>Title: The sodium ion translocating oxaloacetate decarboxylase of Klebsiella pneumon
A:Reference number: A28088; MUID:88257085; PMID:2454915
A:Accession: A28088
A:Molecule type: DNA
A:Residues: 1-596 <SCCH>
A:Cross-references: EMBL:J03885; NID:g149288; PIDN:AAA25120.1; PTD:g149289
C:Superfamily: Klebsiella pneumoniae oxaloacetate decarboxylase alpha chain; lipoYL/B
C:Keywords: biotin binding; carbon-carbon lyase; carboxy-lyase; sodium pump
F:552-596/domain: lipoYL/biotin-binding homology <PDB>
F:562/binding site: biotin (lys) (covalent) #status predicted

Query Match 48.2%; Score 165; DB 2; Length 596;
Best Local Similarity 53.0%; Pred. No. 6, 3e-09;
Matches 35; Conservative 9; Mismatches 22; Indels 0; Gaps 0;

OY 4 IPAPLAGTGVSKILVKEGDFVAKQTVLVLEAMKETETINAPTDRKVEKVLVKERDAVQG 63
:|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 530 VTAPLAGTGVKVLASEGTVAAGVLLILEAMKETETIRAQAQTGVIAKADAVAVG 589

OY 64 QGLIKI 69
|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 590 DTLMFTL 595

RESULT 13
G82308
oxaloacetate decarboxylase, alpha chain VC0550 [similarity] - Vibrio cholerae (strain
C:Species: Vibrio cholerae
C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
C:Accession: G82308
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.
chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Basy, S.; Qin, H.; Dragol, I.; Sellers
l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A>Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A:Reference number: AB2035; MUID:20406833; PMID:10952301
A:Accession: G82308
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-597 <HEI>

A:Cross-references: GB:AE004141; GB:AE003852; NID:96654976; PIDN:AA93718.1; GSPDB:GN003
A:Experimental source: serogroup O1; strain N16561; biotype El Tor
C:Genetics:
A:Gene: VC0550
A:Map position: 1
C:Superfamily: Klebsiella pneumoniae oxalacetate decarboxylase alpha chain; lipoyl/biot

Query Match	48.0%;	Score 164;	DB 2;	Length 597;
Best Local Similarity	50.0%;	Pred. No. 7.9e-09;		
Matches 33;	Conservative 14;	Mismatches 19;	Indels 0;	Gaps 0;

```

QY      4 IPAPLAGTSKILVKEGSDTVKAGQTVLVLKAMKMETELNAPTDGEVKVLYKERDAVQGG 63
      : |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||:
DB      531 VAAPLAGTIFKIQVEGDEVAEGDVLITLAKMKETELRARSQVIGLHLHKEGSDSVRG 590

```

RESULT 14
D81367

C:Probable pyruvate carboxylase (EC 6.4.1.1) B chain Cj0933c [imported] - Campylobacter jejuni
C:Species: Campylobacter jejuni
C:Date: 31-Mar-2000 #sequence.revision 31-Mar-2000 #text.change 03-Jun-2002
C:Accession: DB1367
R:Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chillingworth,
C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; Vanyilet, A.; Whitehead, S.; Barrell, R.
Nature 403, 665-668, 2000
A:title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hyp
#:Reference number: AB1230; MUID:20150912; PMID:10688204

A:Accession: D8136/
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-599 <PAR>
A:Cross-references: GB:AL139076; GB:AL111168; NTID:66968128; PIDN:CAR73190.1; PID:669683
A:Experimental source: serotype O2, strain NCTC 11168
C:Genetics:
A:Gene: pycB; Cj0933c
:Keywords: ligase

[illegible]

OY 61 QGGGGL 66
 | : |
 Db 589 NEGEVL 594

RESULT 15
 D69510
 oxaloacetate decarboxylase, biotin carboxyl carrier subunit homolog - Archaeoglobus fulgidus
 C:Species: Archaeoglobus fulgidus
 C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 20-Aug-1999

C: Superfamily: biotin carboxyl carrier protein; lipoyl/biotin-binding homology
E: 66-139/Domain: lipoyl/biotin-binding homology <LPB>

Query Match	47.1%;	Score 161;	DB 2;	Length 142;
Best Local Similarity	50.8%;	Pred. No. 3.6e-09;		
Matches 32;	Conservative 13;	Mismatches 18;	Indels 0;	Gaps 0;

```

QY      1 EGEIAPPLAGIVSKILYKEGDYTKAGQVYLVLANKMETELNAPITGQVEKVLKERRAV 60
      | : : | | | | | | | | : : : | | | | : | | : : | | | :
Db      70 ENAVISMPLPGVVLKILVKGDKVAGEPVIVESMKMENEIVSPTEGVAAELIVKEGRI 129

```

RESULT 16
H98247
biotin carboxylase protein A2 [imported] - Agrobacterium tumefaciens (strain C58, Cer
C:Species: Agrobacterium tumefaciens
C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 11-Jan-2002
C:Accession: H98247

Query Match	46.68;	Score 159.5;	DB 2;	Length 576;
Best Local Similarity	50.78;	Pred. No. 2.2e-08;		
Matches 35;	Conservative 13;	Mismatches 20;	Indels 1;	Gaps 1

[illegible]

```

RESULT 17
AC3038
biotin carboxylase [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
C.Species: Agrobacterium tumefaciens
C.Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 01-Feb-2002
C.Accession: AC3038
R.Wood, D.W., Setubal, J.C., Kaul, R., Monks, D., Chen, L., Wood, G.E., Chen, Y.
et al., G., Gillet, W., Grant, C., Guenther, D., Kutayavin, T., Levy, R., Li, M
; Karp, P.; Romero, P.; Zhang, S.

```


Query March	44.7%	Score 153;	DB 2;	Length 620;
Best Local Similarity	45.6%;	Pred. No. 1e-07;		
Matches 31;	Conservative 14;	Mismatches 23;	Indels 0;	Gaps 0;

Oy	2	GEIPPLAGTWSKILYKEDDYKACQATVILEAMKRETEINAPTDQKVKVLYKERDAVO	61
Db	545	GDVSPITGKVNKNVNGDEVKEDDYLLVYAMKKNENIHSPPGVGIVIEIPVRGEITN	604

Query Match 44.4%; Score 152; DB 2; Length 436;
Best Local Similarity 51.6%; Pred. No. 9.1e-08;
Matches 32; Conservative 11; Mismatches 19; Indels 0; Gaps

DB 367 GYIAPLAGVSKILVKEGDTVRKAGQTVLVLEAMKMETEINAPDGVKVERDAVGG 426
OY 62 GG 63
DB 427 DG 428

RESULT 26
B44465
sodium ion pump oxaloacetate decarboxylase subunit alpha - Salmonella typhimurium
C:Species: Salmonella typhimurium
C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 18-Aug-2000
C:Accession: B44465
R:Woehli, G.; Wifling, K.; Dimroth, P.
J. Biol. Chem. 267, 22798-22803, 1992
A>Title: Sequence of the sodium ion pump oxaloacetate decarboxylase from Salmonella typhimurium
A:Reference number: A44465; MUID:93054591; PMID:1331067
A:Accession: B44465
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-591 <MOE>
A:Cross-references: GB:M96434; NID:g154195; PIDN:AAA02973.1; PID:g408892
A:Experimental source: IRT2
A>Note: sequence extracted from NCBI backbone
C:Superfamily: Klebsiella pneumoniae oxaloacetate decarboxylase alpha chain; lipoyl/biotin F:518-591/Domain: lipoyl/biotin-binding homology <LPB>
F:557/Binding site: biotin (Lys) (covalent) #status predicted

Query Match
Best Local Similarity 44.4%; Score 152; DB 2; Length 591;
Matches 32; Conservative 9; Mismatches 25; Indels 0; Gaps 0;

OY 4 IPAPLAGVSKILVKEGDTVRKAGQTVLVLEAMKMETEINAPDGVKVERDAVGG 63
DB 525 VYAPLAGNIMKVIATGEGVVAEGDVLILLEAMKMETEINAAQAGTVRGAVKSGDAVSVG 584

OY 64 GGLIKI 69
DB 585 DTLMTL 590

RESULT 27
E83791
hypothetical protein BH1133 [imported] - Bacillus halodurans (strain C-125)
C:Species: Bacillus halodurans
C>Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C:Accession: E83791
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hirai
Nucleic Acids Res. 28, 4317-4331, 2000
A>Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A:Reference number: A83650; MUID:20512582; PMID:11058132
A:Accession: E83791
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-70 <STO>
A:Cross-references: GB:AP001511; GB:BA000004; NID:g10173727; PIDN:BA04852.1; GSPDB:GN00
A:Experimental source: strain C-125
C:Genetics:
A:Gene: BH1133

Query Match
Best Local Similarity 43.9%; Score 150; DB 2; Length 70;
Matches 31; Conservative 15; Mismatches 21; Indels 0; Gaps 0;

OY 3 EIPAPLAGVSKILVKEGDTVRKAGQTVLVLEAMKMETEINAPDGVKVERDAVGG 62
DB 3 KIETNAGNWKVILVKEGQVDEGVATLESKMEIPVBAASGGTVKSVLKQEGEFTDE 62

OY 63 GGLIKI 69
DB 63 GGLIKI 69

RESULT 28
C71667
propionyl-CoA carboxylase alpha chain precursor (pcca) RP618 - Rickettsia prowazekii
C:Species: Rickettsia prowazekii
C>Date: 21-Nov-1998 #sequence_revision 21-Nov-1998 #text_change 03-Nov-2000
C:Accession: C71667
R:Andersson, S.G.E.; Zomrodipour, A.; Andersson, J.O.; Slicheritz-Ponten, T.; Alsmark
Nature 396, 133-140, 1998
A>Title: The genome sequence of Rickettsia prowazekii and the origin of mitochondria.
A:Reference number: A11630; MUID:99039499; PMID:9823893
A:Accession: C71667
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-665 <AND>
A:Cross-references: GB:AJ235272; GB:AJ235269; NID:g3861033; PIDN:CAA15061.1; PID:g386
A:Experimental source: strain Madrid E
C:Genetics:
A:Gene: pcca; RP618
C:Superfamily: propionyl-CoA carboxylase alpha chain; biotin carboxylase homology; 11 F:592-665/Domain: lipoyl/biotin-binding homology <LPB>

Query Match
Best Local Similarity 43.9%; Score 150; DB 2; Length 665;
Matches 31; Conservative 13; Mismatches 22; Indels 0; Gaps 0;

OY 3 EIPAPLAGVSKILVKEGDTVRKAGQTVLVLEAMKMETEINAPDGVKVERDAVGG 62
DB 598 EIQAPLSGQIAIKVKEGQVETIGQIEMILFAMKMETEINLAERDGIKIFVNERDNVVR 657

OY 63 GGLIKI 68
DB 658 GKILDE 663

RESULT 29
AB0509
oxaloacetate decarboxylase (EC 4.1.1.3) alpha chain [imported] - Salmonella enterica
C:Species: Salmonella enterica subsp. enterica serovar Typhi
A>Note: this species has also been called Salmonella typhi
C>Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 27-Nov-2001
C:Accession: AB0509
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Church
th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farr
S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens,
A>Title: Complete genome sequence of a multiple drug resistant Salmonella enterica se
A:Reference number: AB0502; PMID:11677608
A:Accession: AB0509
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-591 <PAR>
A:Cross-references: GB:AL513382; PIDN:CAD01210.1; PID:g16501339; GSPDB:GN00176
C:Genetics:
A:Gene: oadA
C:Superfamily: Klebsiella pneumoniae oxaloacetate decarboxylase alpha chain; lipoyl/b
C:Keywords: carbon-carbon lyase; carboxy-lyase

Query Match
Best Local Similarity 43.3%; Score 148; DB 2; Length 591;
Matches 31; Conservative 10; Mismatches 25; Indels 0; Gaps 0;

OY 4 IPAPLAGVSKILVKEGDTVRKAGQTVLVLEAMKMETEINAPDGVKVERDAVGG 63
DB 525 VYAPLAGNIMKVIATGEGVVAEGDVLILLEAMKMETEINAAQAGTVRGAVKSGDAVSVG 584

OY 64 GGLIKI 69
DB 585 DTLMTL 590

RESULT 30

AEO909
 oxaloacetate decarboxylase (EC 4.1.1.3) alpha chain [Imported] - Salmonella enterica subsp.
 C:Species: Salmonella enterica subsp. enterica serovar Typhi
 A:Note: this species has also been called Salmonella typhi
 C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 27-Nov-2001
 C:Accession: AE0909
 R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher, T.;
 S.: Moule, S.; O'Gaora, P.
 A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; et al.
 A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serovar
 A:Reference number: AB0502; PMID:1167608
 A:Accession: AE0909
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1591 <PAR>
 A:Cross-references: GB:AL513382; PIDN:CAD07867.1; PID:g16504414; GSPDB:GN00176
 C:Genetics:
 A:Gene: oadA
 C:Superfamily: Klebsiella pneumoniae oxaloacetate decarboxylase alpha chain; lipoyl/Blot
 C:Keywords: carbon-carbon lyase; carboxy-lyase

 Query Match 43.3%; Score 148; DB 2; Length 591;
 Best Local Similarity 47.0%; Pred. NO. 3.2e-07;
 Matches 31; Conservative 10; Mismatches 25; Indels 0; Gaps 0;

 Oy 4 IPAPLAGTIVSKLIVEGGTVKAGCTVLEAMKMETETINAPDGEVEVLYKENDAYOGG 63
 Db 525 VTAPlAGNIMKVIATIEGGSVAEGDVLLILEAMKMETETIRAAQAGTVRGIAVKSGDVAVSG 584
 Oy 64 GGLIKI 69
 Db 585 DTLMTL 590

 Search completed: May 1, 2003, 07:50:28
 Job time : 11.5729 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 1, 2003, 07:49:49 ; Search time 6.19792 Seconds

(without alignments)
468.439 Million cell updates/sec

Title: US-09-987-485-2

Perfect score: 342
Sequence: 1 EEEIPAPLAGTVSKILVKEG.....KYLVERDAVGQGLIKIG 70

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	342	100.0	123	1	BCCP_PROER
2	173	50.6	567	1	PCYB_METUA
3	165	48.2	595	1	DCOA_KLEPN
4	152	44.4	590	1	DCOA_SALTY
5	148	42.3	717	1	BCCA_MOUSE
6	145	42.4	130	1	BCCP_STRTMU
7	143	41.8	70	1	BCCP_LYCES
8	143	41.8	725	1	MCCA_HUMAN
9	143	41.8	734	1	MCCA_ARATH
10	140	40.9	568	1	PCYB_METTH
11	138.5	40.5	157	1	BCCP_PORPU
12	138	40.4	1180	1	PCY2_YEAST
13	133	38.9	1178	1	PCY1_YEAST
14	132	38.6	1189	1	PCY_PICPA
15	130	38.0	598	1	BCCA_MYCLE
16	130	38.0	731	1	MCCA_SOYBN
17	129	37.7	1178	1	PCY_MOUSE
18	126.5	37.0	181	1	BCCP_ANASP
19	124.5	36.4	654	1	BCCA_MYCTU
20	122	35.7	164	1	BCCP_CHITR
21	121	35.4	1178	1	PCY_HUMAN
22	119	34.8	1178	1	PCY_RAT
23	118	34.5	70	1	BTRB_MYCSM
24	118	34.5	163	1	BCCP_CHLMU
25	115.5	33.8	547	1	ODP2_PSEAE
26	115	33.6	70	1	BTRB_MYCTU
27	114	33.3	438	1	ODP2_MYCA
28	113.5	33.2	629	1	ODP2_ECOLI
29	111	32.5	704	1	PCCA_RAT
30	110	32.2	703	1	PCCA_HUMAN
31	108.5	31.7	167	1	BCCP_CHLPN
32	108	31.6	553	1	ODP2_MYCTU
33	107	31.3	70	1	BTRB_MYCLE

34	104	30.4	1835	1	DUR1_YEAST	P32528 saccharomyc
35	103.5	30.3	567	1	ODP2_HAEIN	P45118 haemophilus
36	103	30.1	462	1	ODPB_ZYMO	P06113 zymomonas m
37	102.5	30.0	155	1	BCCP_HAEIN	P43874 haemophilus
38	102	29.8	544	1	ODP2_ACHLA	P35489 azotobacter
39	102	29.8	637	1	ODP2_AZOVI	P10802 azotobacter
40	101.5	29.7	262	1	BCCP_SOYBN	P42783 glycine max
41	98.5	28.8	156	1	BCCP_PSEAE	P32799 pseudomonas
42	97.5	28.5	156	1	BCCP_ECOLI	P02905 escherichia
43	97	28.4	384	1	ODP2_MGE	P47514 mycoplasma
44	95	27.8	463	1	ODP2_YEAST	P19262 saccharomyc
45	94.5	27.6	152	1	BCCP_YEAST	P19262 saccharomyc
46	94	27.5	447	1	ODP2_YEAST	P19262 saccharomyc
47	93.5	27.3	2273	1	HEAL_YEAST	P32874 saccharomyc
48	93	27.2	402	1	ODP2_MYCPN	P75392 mycoplasma
49	92	27.2	440	1	ODP2_ZYMO	P06119 zymomonas m
50	92	26.9	427	1	ODP2_BACST	P11961 bacillus st
51	91	26.6	398	1	ACOC_BACSU	P31550 bacillus st
52	91	26.6	401	1	ODP2_RICPR	P92474 rickettsia
53	91	26.6	417	1	ODP2_BACSU	P16283 bacillus su
54	88	25.7	430	1	ODP2_STRAU	P59881 staphylococ
55	86.5	25.3	553	1	ODP2_ALCEU	P05098 alcaligenes
56	85.5	25.0	159	1	BCCP_BACSU	P49786 bacillus su
57	84.5	24.7	280	1	BCCP_ARATH	P42533 arabidopsis
58	84	24.6	460	1	ODPB_RHIME	P09194 rhizobium m
59	82	24.0	501	1	ODP2_HUMAN	P00330 homo sapien
60	81.5	23.8	420	1	ODP2_BUCAT	P57389 buchnera ap
61	79.5	23.2	398	1	ODP2_AZOVI	P20708 azotobacter
62	79	23.1	409	1	ODP2_FUGRU	P90512 fuugu rubrip
63	79	23.1	2280	1	COAC_SCHPO	P78820 schizosach
64	78	22.8	453	1	ODP2_HUMAN	P36957 homo sapien
65	77.5	22.8	455	1	ODP2_PIG	P09011 sus scrofa
66	77.5	22.7	409	1	ODP2_HAEIN	P45302 haemophilus
67	77.5	22.7	424	1	ODP2_BACSU	P37922 bacillus su
68	77.5	22.7	441	1	ODP2_BACSU	P21883 bacillus su
69	77.5	22.7	441	1	ODP2_SCHPO	P04681 schizosach
70	77	22.5	2483	1	COA2_HUMAN	P000763 homo sapien
71	76.5	22.4	819	1	RNFC_HAEIN	P71397 haemophilus
72	74.5	21.8	416	1	ODP2_ALCEU	P52993 alcaligenes
73	74	21.6	408	1	ODP2_RICPR	P92470 rickettsia
74	73	21.3	482	1	ODP2_BOVIN	P11811 bos taurus
75	72	21.1	442	1	ODP2_RAT	P001205 rattus norv
76	72	21.1	482	1	ODP2_HUMAN	P11812 homo sapien
77	71.5	20.9	355	1	YHII_ECOLI	P37626 escherichia
78	71	20.8	370	1	ACOC_PSEPU	P59655 pseudomonas
79	70.5	20.6	329	1	AN36_HELPY	P94851 escherichia
80	70.5	20.6	541	1	YOE8_ECOLI	P46808 escherichia
81	70	20.5	336	1	ODP2_BUCAT	P57302 buchnera ap
82	70	20.5	448	1	PRTE_ERWCH	P23597 erwina chr
83	70	20.5	482	1	ODP2_YEAST	P12695 saccharomyc
84	69	20.2	180	1	GCSH_DEIRA	P09173 delnoccocus
85	69	20.2	482	1	ODP2_MOUSE	P53395 mus musculu
86	68.5	20.0	329	1	AN36_HELPY	P94851 escherichia
87	68	19.9	131	1	GCSH_HELPY	P09341 heliocobacte
88	68	19.9	163	1	GCSH_XYMPA	P09347 xylella fas
89	68	19.9	410	1	GCSH_MESCR	P93255 mesobryant
90	67.5	19.7	592	1	ODP2_YEAST	P16451 saccharomyc
91	67	19.6	124	1	GCSH_DICDI	P36413 dictyosteli
92	67	19.6	373	1	ACOC_ALCEU	P09143 thermoplasma
93	67	19.6	614	1	ODP2_HUMAN	P27747 alcaligenes
94	67	19.6	2346	1	COAL_BOVIN	P10515 h dihydrotol
95	67	19.6	2346	1	COAL_SHEEP	P09143 thermoplasma
96	66.5	19.4	1524	1	RPOC_THEDA	P02859 ovis aries
97	66	19.3	432	1	APRE_PSEAE	P09143 thermoplasma
98	66	19.3	2233	1	COAC_YEAST	P003025 pseudomonas
99	66	19.3	2345	1	COAL_RAT	P00095 saccharomyc
100	66	19.3	2346	1	COAL_HUMAN	P11437 rattus norv

ALIGNMENTS

RESULT 1

BCCP_PROF

ID	BCCP_PROF	STANDARD:	PRT:	123 AA.
AC	P02904;			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	21-JUL-1986 (Rel. 01, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Biotin carboxyl carrier protein of methylmalonyl-CoA carboxyl- transferase (transcarboxylase, 1.3S subunit).			
OS	Propionibacterium freudenreichii shermanii.			
OC	Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;			
OC	Actinomycetales; Propionibacteriaceae; Propionibacteriaceae;			
OC	Propionibacterium.			
OX	NCBI_TaxID=1752;			
RN	[1]			
RN	SEQUENCE FROM N.A.			
RA	MEDLINE=85298212; PubMed=3898065;			
RA	Murtif V.L., Bahler C.R., Samols D.;			
RT	"Cloning and expression of the 1.3S biotin-containing subunit of transcarboxylase."			
RT	Proc. Natl. Acad. Sci. U.S.A. 82:5617-5621(1985);			
RL	[2]			
RN	SEQUENCE.			
RA	MEDLINE=80049796; PubMed=40985;			
RA	Maloy W.L., Bowlen B.U., Zwolinski G.K., Kumar G.K., Wood H.G.,			
RA	Ericsson L.H., Walsh K.A.;			
RT	"Amino-acid sequence of the biotinyl subunit from transcarboxylase";			
RT	J. Biol. Chem. 254:11615-11622(1979);			
RL	[3]			
RN	MUTAGENESIS OF ALA-87; MET-88; LYS-89 AND MET-90.			
RA	MEDLINE=97406744; PubMed=1526981;			
RA	Shenoy B.C., Xie Y., Park V.L., Kumar G.K., Beegen H., Wood H.G.,			
RA	Samols D.;			
RT	"The importance of methionine residues for the catalysis of the biotin enzyme, transcarboxylase. Analysis by site-directed mutagenesis."			
RT	J. Biol. Chem. 267:18407-18412(1992);			
RL	[4]			
RN	STRUCTURE BY NMR.			
RA	MEDLINE=98066416; PubMed=9398186;			
RA	Reddy D.V., Shenoy B.C., Carey P.R., Soenichsen F.D.;			
RT	"Absence of observable biotin-protein interactions in the 1.3S subunit of transcarboxylase: an NMR study."			
RT	Biochemistry 36:14676-14682(1997);			
RL	[5]			
RN	FUNCTION: THE BIOTINYL 1.3S SUBUNIT SERVES AS A CARBOXYL CARRIER BETWEEN THE SUBSTRATE BINDING SITES ON THE 12S AND 5S SUBUNITS.			
CC	-1- SUBUNIT: TRANSCARBOXYLASE IS COMPOSED OF THREE SUBUNITS: 1.3S, 5S, AND 12S. THE CORE OF THE ENZYME IS COMPOSED OF SIX 12S SUBUNITS. ON EACH SIDE OF THE CORE THERE ARE THREE PAIRS OF 5S SUBUNITS. EACH 5S DIMER IS ATTACHED TO THE CORE BY TWO 1.3S SUBUNITS. THUS THE TOTAL NUMBER OF CHAINS IS 30 (6 + 12 + 12).			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL: M1738; AAA25674.1; .			
DR	PIR: A03401; BKIP.			
DR	HSSP: P02905; 1A6X.			
DR	InterPro: IPR001882; Biotin-attach.			
DR	InterPro: IPR000089; Biotin_lipoyl.			
DR	Pfam: PF00364; biotin_lipoyl.1.			
DR	PROSITE: PS00186; BIOTIN; 1.			
FT	Biotin.			
KW	BIOTIN.			
FT	BINDING 89 89 BIOTIN.			
SQ	SEQUENCE 123 AA; 12367 MW; D0980C2065EA9A89 CRC64;			
Query Match	100.0%; Score 342; DB 1; Length 123;			
Best Local Similarity	100.0%; Pred. No. 1; le-27;			
Matches	70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			

QY 1 EGEIPAPLAGVSKILVEGDTVKNAGTGVLEAKMKETINAPDGKVEKVEKRDV 60
DB 54 EGEIPAPLAGVSKILVEGDTVKNAGTGVLEAKMKETINAPDGKVEKVEKRDV 113
QY 61 OGCGGLIRIG 70
DB 114 OGCGGLIRIG 123

RESULT 2

ID	PCYC_METUA	STANDARD:	PRT:	567 AA.
AC	Q58628;			
DT	30-MAY-2000 (Rel. 39, Created)			
DT	30-MAY-2000 (Rel. 39, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Pyruvate carboxylase subunit B (EC 6.4.1.1) (pyruvic carboxylase B).			
OS	PCB OR M01231.			
OS	Methanococcus jannaschii.			
OC	Archaea; Euryarchaeota; Methanococci; Methanococcales;			
OC	Methanocaldococcaceae; Methanocaldococcus.			
OX	NCBI_TaxID=2190;			
RN	[1]			
RN	SEQUENCE FROM N.A.			
RA	STRAIN=JAL-1 / DSM 2661 / ATCC 43067;			
RA	MEDLINE=96337999; PubMed=8688087;			
RA	Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,			
RA	Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,			
RA	Kerlavage A.R., Dougherty B.A., Tomb J.F., Adams M.D., Reich C.I.,			
RA	Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,			
RA	Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhrman J.L., Nguyen D.,			
RA	Uitterlinden T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,			
RA	Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,			
RA	Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;			
RT	"Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii."			
RT	Science 273:1058-1073(1996);			
RL	[2]			
RN	SEQUENCE OF 190-125; 260-270; 277-289; 309-325; 328-358;			
RA	370-380; 386-409; 422-438; 491-506 AND 491-506, AND FUNCTION.			
RA	MEDLINE=21034791; PubMed=11195096;			
RA	Mukhopadhyay B., Patel V.J., Wolfe R.S.;			
RT	"A stable archaeal pyruvate carboxylase from the hyperthermophile Methanococcus jannaschii."			
RT	Arch. Microbiol. 174:406-414(2000);			
CC	-1- FUNCTION: PYRUVATE CARBOXYLASE CATALYZES A 2-STEP REACTION, INVOLVING THE ATP-DEPENDENT CARBOXYLATION OF THE COVALENTLY ATTACHED BIOTIN IN THE FIRST STEP AND THE TRANSFER OF THE CARBOXYL GROUP TO PYRUVATE IN THE SECOND.			
CC	-1- CATALYTIC ACTIVITY: ATP + pyruvate + HCO(3)(-) = ADP + phosphate + oxaloacetate.			
CC	-1- COFACTOR: ATP, MAGNESIUM (OR MANGANESE OR COBALT), PYRUVATE AND BICARBONATE.			
CC	-1- ENZYME REGULATION: INHIBITED BY MAGNESIUM, WHEN ITS CONCENTRATION EXCEEDED THE ATP ONE, AND BY HIGH CONCENTRATION OF ATP AND ALPHA- KETOGLOTERATE.			
CC	-1- PATHWAY: GLUCONEOGENESIS.			
CC	-1- SUBUNIT: HETEROCTAMER OF FOUR A AND FOUR B SUBUNITS.			
CC	-1- MASS SPECTROMETRY: MM-64160; METHOD-MALDI.			
CC	-1- MISCELLANEOUS: ITS OPTIMUM PH IS 8.5 AND THE OPTIMUM TEMPERATURE IS 80-90 DEGREES CELSIUS.			
CC	-1- SIMILARITY: WITH OTHER BIOTIN CARBOXYLASES, LIPAMIDE TRANSFERASES AND CARBAMYL PHOSPHATE SYNTHETASES.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).			
CC	-----			

DR EMBL: U67563; AAB99233.1; -
DR HSSP: P02905; IAXX.
DR TIGR: M1231; -
DR InterPro: IPR001882; Biotin_attach.
DR InterPro: IPR000089; Biotin_lipoyl.
DR InterPro: IPR000891; HMG1-like.
DR InterPro: IPR003379; PTC_OADA.
DR Pfam: PF00364; biotin_lipoyl.1.
DR Pfam: PF00682; HMG1-like.1.
DR Pfam: PF02436; PTC_OADA.1.
DR TIGRfams: TIGR01108; oada.1.
DR PROSITE: PS00188; BIOTIN.1.
KM Ligase: Multifunctional enzyme; Gluconeogenesis; Magnesium; Pyruvate;
FT Biotin; Complete proteome.
KW BINDING 52 60 PYRUVATE (BY SIMILARITY).
FT BINDING 533 533 BIOTIN (BY SIMILARITY).
SQ SEQUENCE 567 AA: 63907 MW: 5E07800622545628 CRC64:

Query Match 50.6%; Score 173; DB 1; Length 567;
Best Local Similarity 50.7%; Pred. No. 4e-10;
Matches 35; Conservative 12; Mismatches 22; Indels 0; Gaps 0;

QY 1 EGGIPAPLAGTYSKIIYKEDGVYKAGQVLYLEAMKETEINAPFDGKYEVLYKRDVAV 60
DB 498 EGAIVTSPPFGMWYKIKVKEGDKVYKGVIVLEAMKMEHPESVETVERILLIDEGDAV 557
QY 61 QGGGGLIKI 69
DB 558 NVGDVIMII 566

RESULT 3
DCOA_KLEPN STANDARD; PRT: 595 AA.
ID DCOA_KLEPN
AC P13187;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 01-JUL-1993 (Rel. 26, Last annotation update)
DE Oxaloacetate decarboxylase alpha chain (EC 4.1.1.3).
GN OADA.
OS Klebsiella pneumoniae.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Klebsiella.
OX NCBI_Taxid=573;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88257085; PubMed=2454915;
RA Schwarz E., Oesterheld D., Reinke H., Beyreuther K., Dimroth P.;
RT The sodium ion translocating oxaloacetate decarboxylase of Klebsiella
RT pneumoniae. Sequence of the biotin-containing alpha-subunit and
RT relationship to other biotin-containing enzymes.";
RL J. Biol. Chem. 263:9640-9645(1988).
CC -1- FUNCTION: LYASE AND SODIUM TRANSPORTER.
CC -1- CATALYTIC ACTIVITY: Oxaloacetate - pyruvate + CO(2).
CC -1- COFACTOR: BIOTIN AND REQUIRES A SODIUM ION.
CC -1- SUBUNIT: COMPOSED OF THREE CHAINS (ALPHA, BETA, AND GAMMA).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: J03885; AAA25120.1; -
DR PIR: A28088; A28088.
DR HSSP: P02905; IAXX.
DR InterPro: IPR001882; Biotin_attach.
DR InterPro: IPR000089; Biotin_lipoyl.
DR InterPro: IPR000891; HMG1-like.
DR InterPro: IPR003379; PTC_OADA.
DR Pfam: PF00364; biotin_lipoyl.1.
DR HSSP: P02905; IAXX.

DR Pfam: PF00682; HMG1-like.1.
DR Pfam: PF02436; PTC_OADA.1.
DR TIGRfams: TIGR01108; oada.1.
DR PROSITE: PS00188; BIOTIN.1.
KM Decarboxylase: Lyase; Sodium transport; Biotin.
FT INIT_MET 0 0
FT BINDING 561 561 BIOTIN (BY SIMILARITY).
SQ SEQUENCE 595 AA: 63402 MW: AA744A95A6E9488C CRC64:

Query Match 48.2%; Score 165; DB 1; Length 595;
Best Local Similarity 53.0%; Pred. No. 2.7e-09;
Matches 35; Conservative 9; Mismatches 22; Indels 0; Gaps 0;

QY 4 IPAPLAGTYSKIIYKEDGVYKAGQVLYLEAMKETEINAPFDGKYEVLYKRDVAV 63
DB 529 VTPAPLAGTITMKVLAISGQVYAAEGVLLILEAMKETEINAPAGQVYRGVAVAGDAVAVG 588
QY 64 QGLIKI 69
DB 589 DFLMTL 594

RESULT 4
DCOA_SALTY STANDARD; PRT: 590 AA.
ID DCOA_SALTY
AC Q03030;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Oxaloacetate decarboxylase alpha chain (EC 4.1.1.3).
GN (OADA1 OR OADA OR STM0055) AND (OADA2 OR STM3352).
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmonella.
OX NCBI_Taxid=602;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93054591; PubMed=1331067;
RA Moenke G., Wiffling K., Dimroth P.;
RT *Sequence of the sodium ion pump oxaloacetate decarboxylase from
RT Salmonella typhimurium.";
RL J. Biol. Chem. 267:22798-22803(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=LT2 / SGSC1412 / ATCC 700720;
RX MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RT *Complete genome sequence of Salmonella enterica serovar Typhimurium
RT LT2.";
RL Nature 413:852-856(2001).
CC -1- FUNCTION: LYASE AND SODIUM TRANSPORTER.
CC -1- CATALYTIC ACTIVITY: Oxaloacetate - pyruvate + CO(2).
CC -1- COFACTOR: BIOTIN AND REQUIRES A SODIUM ION.
CC -1- SUBUNIT: COMPOSED OF THREE CHAINS (ALPHA, BETA, AND GAMMA).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: M96434; AAA02973.1; -
DR DR AEO08696; AAL19019.1; -
DR EMBL: AEO08696; AAL19019.1; -
DR PIR: B44465; B44465.
DR HSSP: P02905; IAXX.

```

DR StyGene; SG10259; oada1.
DR StyGene; SG77777; oada2.
DR InterPro; IPR001882; Biotin_attach.
DR InterPro; IPR000089; Biotin_1lpoyl.
DR InterPro; IPR000891; HMGL-1like.
DR InterPro; IPR003379; PYC_OADA.
DR Pfam; PF00364; biotin_1lpoyl; 1.
DR Pfam; PF00682; HMGL-1like; 1.
DR Pfam; PF02436; PYC_OADA; 1.
DR TIGRfams; TIGR01108; oada; 1.
DR PROSITE; PS00188; BIOTIN; 1.
KW Decarboxylase; Lyase; Sodium transport; Biotin; Complete proteome.
FT INT_MPT 0
FT BINDING 556 556 BIOTIN (By SIMILARITY).
SQ SEQUENCE 590 AA: 63075 MW: 4E421F9324AED7B CRC64:

Query Match 44.4%; Score 152; DB 1; Length 590;
Best Local Similarity 48.5%; Pred. No. 5.3e-08;
Matches 32; Conservative 9; Mismatches 25; Indels 0; Gaps 0;

OY 4 IPAPLAGVSKILVEGDTVAGQTVLVLEAKMETETINAPTDGKVEKYLVERDVG 63
D6 534 VTAPLAGNWKVIAEGQTVAGDVLLILEAKMETETIRAAQAGTVRGIAVSGDAVS 583
OY 64 GGLIKI 69
D6 584 DTLMTL 589

RESULT 5
MCCA_MOUSE STANDARD; PRT; 717 AA.
AC 099NR8; Q9D8R2;
AT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Methylcrotonyl-CoA carboxylase alpha chain, mitochondrial precursor
DE (EC 6.4.1.4) (3-methylcrotonyl-CoA carboxylase 1) (MCCase alpha
DE subunit) (3-methylcrotonyl-CoA:carbon dioxide ligase alpha subunit).
GN MCCI OR MCCA.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN NM
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6;
RC MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Aikawa T., Hara A., Nishikishi Y., Konno H., Adachi J., Fukuda S.,
RA Iizawa K., Iwawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaoka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batilov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochava H.,
RA Kuehl P., Lewis S., Matsuo Y., Nakai T., Pesole G., Quackenbush J.,
RA Schiraldi L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okitani T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carolini P., de Bonaldi M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Guclincich S., Hill D., Hoffmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaserts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,

```

```

RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kotsuki S.,
RA Hayashizaki Y.,
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RA Strausberg R.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: ATP + 3-methylcrotonyl-CoA + HCO(3)(-) = ADP +
CC phosphate + 3-methylglutaconyl-CoA.
CC -1- COFACTOR: Biotin.
CC -1- PATHWAY: Leucine catabolism.
CC -1- SUBUNIT: Probably a dodecamer composed of six biotin-containing
CC alpha subunits and six beta subunits.
CC -1- SUBCELLULAR LOCATION: Mitochondrial matrix (By similarity).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
D6 EMBL; AF310338; AAG50244.1; -
D6 EMBL; AK007782; BAB25253.1; -
D6 EMBL; BC021382; AAH21382.1; -
D6 HSSP; P24182; IDV1.
D6 MGD; MGI:1919289; Mcccl.
DR InterPro; IPR001882; Biotin_attach.
DR InterPro; IPR000089; Biotin_1lpoyl.
DR InterPro; IPR000901; CPSase.
DR Pfam; PF00289; CPSase_L_chain; 1.
DR Pfam; PF00364; biotin_1lpoyl; 1.
DR Pfam; PF02785; biotin_carp_C; 1.
DR Pfam; PF02786; CPSase_L_D2; 1.
DR PROSITE; PS00188; BIOTIN; 1.
DR PROSITE; PS00867; CPSASE_2; 1.
DR MitoChondrion; Ligase; Biotin; ATP-binding; Transalt peptide.
FT TRANSIT 1 44 MITOCHONDRION (POTENTIAL).
FT CHAIN 45 717 METHYLCROTONYL-CoA CARBOXYLASE ALPHA
FT NP_BIND 205 210 ATP (POTENTIAL).
FT ACT_SITE 335 335 BIOTIN (BY SIMILARITY).
FT BINDING 677 677 BIOTIN (BY SIMILARITY).
FT DOMAIN 709 714 POLY-GLU.
FT CONFLICT 324 324 R -> K (IN REF. 3).
FT CONFLICT 507 507 A -> P (IN REF. 1).
SQ SEQUENCE 717 AA: 79343 MW: F653FE7AC1E5AA90 CRC64:

Query Match 43.3%; Score 148; DB 1; Length 717;
Best Local Similarity 50.0%; Pred. No. 1.6e-07;
Matches 28; Conservative 11; Mismatches 17; Indels 0; Gaps 0;

OY 1 EGEIPAPLAGVSKILVEGDTVAGQTVLVLEAKMETETINAPTDGKVEKYLVE 56
D6 642 OGCTIAPMTGTIERVYVAGDVRVAGDSLMIWIAKMEHTIKAPDKRIKVFES 697

RESULT 6
BCCP_STRMU STANDARD; PRT; 130 AA.
AC P29337;
AT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Biotin carboxyl carrier protein (BCCP).
OS Streptococcus mutans.
OC Bacteriia; Firmicutes; Lactobacillales; Streptococcaceae;
OX Streptococcus.
RN NCBI_TaxID=1309;

```

RP SEQUENCE FROM N.A.
RC STRAIN-OT-041 / Serotype C;
RX MEDLINE-93159778; PubMed-8431283;
RA Wang D., Waye M.M., Taricani M., Buckingham K., Sandham H.J.;
RT "Biotin-containing protein as a cause of false positive clones in
RL gene probing with streptavidin/biotin.";
CC Biotechniques 14:209-212(1993).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: M80523; AA003702.1; -
DR HSSP: P02905; IABX.
DR InterPro: IPR001882; Biotin_attach.
DR InterPro: IPR000089; Biotin_lipoyl.
DR Pfam: PF00364; Biotin_lipoyl; 1.
DR PROSITE: PS00188; BIOTIN; 1.
KW Fatty acid biosynthesis; Biotin.
FT BINDING 96 BIOTIN (BY SIMILARITY).
SQ SEQUENCE 130 AA; 13601 MM; A0D6025EC46FF00B CRC64;
Query Match 42.4%; Score 145; DB 1; Length 130;
Best Local Similarity 47.0%; Pred. No. 5.8e-08;
Matches 31; Conservative 8; Mismatches 27; Indels 0; Gaps 0;
QY 4 ITPAPLAGYVSKILYKGGDTYKAGQTVLVLEAMKMEIINAPDGGKVKYKERDAVGG 63
DB 64 MPAPMGITLKLVLNVGDYVSENPMLILEAMKMEIIVAGMGTVAIHVSSGQGVADG 123
QY 64 GGLIKI 69
DB 124 DNLITL 129
RESULT 7
BCCP_LYCES
ID BCCP_LYCES STANDARD; PRT; 70 AA.
AC P05115;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Biotin carboxyl carrier protein of acetyl-CoA carboxylase (BCCP)
DE (Fragment)
OS Lycopersicon esculentum (Tomato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4081;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-87231088; PubMed-3588314;
RA Hoffman N.E., Pichersky E., Cashmore A.R.;
RT "A tomato cDNA encoding a biotin-binding protein.";
RL Nucleic Acids Res. 15:3928-3928(1987).
CC -1- FUNCTION: THIS PROTEIN IS A COMPONENT OF THE ACETYL COENZYME A
CC CARBOXYLASE COMPLEX. FIRST, BIOTIN CARBOXYLASE CATALYZES THE
CC CARBOXYLATION OF THE CARRIER PROTEIN AND THEN THE TRANSCARBOXYLASE
CC TRANSFERS THE CARBOXYL GROUP TO FORM MALONYL-COA.
CC -1- PARTIAL: Long-chain fatty acid biosynthesis; first step.
CC -1- SUBCELLULAR LOCATION: Chloroplast.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC

CC
DR EMBL: Y00144; CAA68339.1; -
DR PIR: A29271; A29271.
DR HSSP: P02905; 1BDO.
DR InterPro: IPR001882; Biotin_attach.
DR InterPro: IPR000089; Biotin_lipoyl.
DR Pfam: PF00364; Biotin_lipoyl; 1.
DR PROSITE: PS00188; BIOTIN; 1.
KW Fatty acid biosynthesis; Biotin; Chloroplast.
FT NON_TER 1
FT BINDING 35 BIOTIN (BY SIMILARITY).
SQ SEQUENCE 70 AA; 7416 MM; D75D018C0BD016BC CRC64;
Query Match 41.8%; Score 143; DB 1; Length 70;
Best Local Similarity 48.4%; Pred. No. 4.9e-08;
Matches 30; Conservative 11; Mismatches 21; Indels 0; Gaps 0;
QY 2 GETPAPLAGYVSKILYKGGDTYKAGQTVLVLEAMKMEIINAPDGGKVKYKERDAV 61
DB 1 GTVAAPVWGLVAVLVKDGKVOGQPVLVLEAMKMEHVAKAPANGVSGLEIKVGSVQ 60
QY 62 GG 63
DB 61 DG 62
RESULT 8
MCCA_HUMAN
ID MCCA_HUMAN STANDARD; PRT; 725 AA.
AC Q96R03; Q9H959; Q9NS97;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE methylcrotonyl-CoA carboxylase alpha chain, mitochondrial precursor
DE (EC 6.4.1.4) (3-Methylcrotonyl-CoA carboxylase 1) (MCCase alpha
DE subunit) (3-methylcrotonyl-CoA:carbon dioxide ligase alpha subunit).
GN MCC1 OR MCCA.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND VARIANTS MCGI ARG-325 AND SER-385.
RX PubMed-11170888;
RA Gallardo M.E., Desviat L.R., Rodriguez J.M., Esparza-Gordillo J.,
RA Perez-Cerda C., Perez B., Rodriguez-Pombo P., Criado O., Sanz R.,
RA Morton D.H., Gibson K.M., Le T.P., Ribes A., Rodriguez de Cordoba S.,
RA Ugarte M., Penalba M.A.;
RT "The molecular basis of 3-methylcrotonylglycinuria, a disorder of
RT leucine catabolism.";
RL Am. J. Hum. Genet. 68:334-346(2001).
RN [2]
RP SEQUENCE FROM N.A., AND VARIANT HIS-464.
RX MEDLINE-21295033; PubMed-11401427;
RA Ooba K., Fukuda T., Morishita R., Abe S., Asakawa S., Yamaguchi S.,
RA Yoshino M., Ihara K., Murayama K., Shigemoto K., Shimizu N., Kondo I.;
RT "Human biotin-containing subunit of 3-methylcrotonyl-CoA carboxylase
RT gene (MCCA): cDNA sequence, genomic organization, localization to
RT chromosomal band 3q27, and expression.";
RL Genomics 72:145-152(2001).
RN [3]
RP SEQUENCE FROM N.A., AND VARIANT MCGI PHE-535.
RX MEDLINE-21299419; PubMed-11406611;
RA Holzinger A., Roesslinger W., Lagler F., Mayerhofer P.U., Lichtner P.,
RA Kattenfeld T., Thny L.P., Nyhan W.L., Koch H.G., Muntau A.C.,
RA Roscher A.A.;
RT "Cloning of the human MCCA and MCCB genes and mutations therein reveal
RT the molecular cause of 3-methylcrotonyl-CoA: carboxylase
RT deficiency.";
RL Hum. Mol. Genet. 10:1299-1306(2001).
RN [4]
RP SEQUENCE FROM N.A., AND VARIANTS MCGI VAL-289; SER-385; PRO-437 AND
RP HIS-532.

RT SSP consortium (Salik/Stanford/REGC).
 RL Submitted (Dec-2001) to the EMBL/Genbank/DBJ databases.
 RN [4]
 RP TISSUE SPECIFICITY.
 RC STRAIN-cv. Landsberg erecta, and cv. Columbia;
 RX MEDLINE-20148760; PubMed-10681539;
 RA McKean A.L., Ke J., Song J., Che P., Achenbach S., Nikolau B.J.,
 RA Wurtele E.S.;
 RT "Molecular characterization of the non-biotin-containing subunit of
 RT 3-methylcrotonyl-CoA carboxylase."
 RL J. Biol. Chem. 275:5582-5590(2000).
 CC -1 CATALYTIC ACTIVITY: ATP + 3-methylcrotonyl-CoA + HCO(3)(-) - ADP +
 CC phosphate + 3-methylglutacetyl-CoA.
 CC -1 COFACTOR: Biotin.
 CC -1 PATHWAY: Leucine catabolism.
 CC -1 SUBUNIT: Probably a heterodimer composed of biotin-containing
 CC alpha subunits and beta subunits (By similarity).
 CC -1 SUBCELLULAR LOCATION: Mitochondrial matrix.
 CC -1 TISSUE SPECIFICITY: In roots, cotyledons, leaves, flowers,
 CC ovaries, siliques and embryos.
 CC -1 MISCELLANEOUS: Temporal and spatial accumulation of the alpha and
 CC beta subunits during development at approximately equal molar
 CC ratios.
 CC -1 CAUTION: Ref.2 sequence differs from that shown due to erroneous
 CC gene model prediction.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U12536; AAA62586.1; -
 DR EMBL: AC006550; AAD25800.1; ALT_SEQ.
 DR EMBL: AF070723; AAL50065.1; -
 DR HSSP: P24182; IBNC.
 DR InterPro: IPR001882; Biotin_attach.
 DR InterPro: IPR000089; Biotin_lipoyl.
 DR InterPro: IPR000901; CPSase.
 DR Pfam: PF02785; Biotin_carb_C.1.
 DR Pfam: PF00364; biotin_lipoyl.1.
 DR Pfam: PF00289; CPSase_L_D2.2.
 DR Pfam: PF02786; CPSase_L_D2.2.
 DR PROSITE: PS00188; BIOTIN.1.
 DR PROSITE: PS00866; CPSASE.1.1.
 DR PROSITE: PS00867; CPSASE.2.1.
 DR Mitochondrion; Ligase; Biotin; ATP-binding; Transl. peptide.
 KW TRANSIT
 FT CHAIN 1 25
 FT CHAIN 26 734
 FT CHAIN 734 734
 FT NP_BIND 198 203
 FT ACT_SITE 329 329
 FT BINDING 699 699
 FT CONFLICT 85 85
 FT CONFLICT 92 92
 FT CONFLICT 281 300
 FT CONFLICT 430 430
 SO SEQUENCE 734 AA; 80451 MW; 251CACF6464B046B CRC64;
 Query Match 41.8%; Score 143; DB 1; Length 734;
 Best Local Similarity 45.6%; Pred. No. 5.3e-07;
 Matches 31; Conservative 10; Mismatches 27; Indels 0; Gaps 0;
 QY 2 GETPAPLAGVSKILVKGSTVAGQVLYLAKMETETINAPDQVEVLYKERAYQ 61
 DB 665 GTTAPAGAGLVAVLVENAKVQGGPILVLEAKMEHVAKDSSGSIQDLKYKAGQOVS 724
 QY 62 GGQGLIKI 69
 DB 725 DGSALFRI 732

RESULT 10
 ID PYCB_METHH STANDARD: PRT; 568 AA.
 AC 027179;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Pyruvate carboxylase subunit B (Ec 6.4.1.1) (Pyruvic carboxylase B).
 GN PYCB OR MTH107.
 OS Methanobacterium thermoautotrophicum.
 OC Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;
 OC Methanobacteriaceae; Methanobacter.
 OX NCBI_TaxID=187420;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Delta H:
 RX MEDLINE-98037514; PubMed-9371463;
 RA Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,
 RA Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert R.,
 RA Harrison D., Hoang L., Keagle P., Lumm W., Pochler B., Qiu D.,
 RA Spadefora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,
 RA Jiwani N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,
 RA McDougall S., Shiner G., Goyal A., Petrovski S., Church G.M.,
 RA Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;
 RT "Complete genome sequence of Methanobacterium thermoautotrophicum
 RT deltaH: functional analysis and comparative genomics."
 RL J. Bacteriol. 179:7135-7155(1997).
 RN [2]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-20.
 RC STRAIN-Delta H:
 RX MEDLINE-98148063; PubMed-9478969;
 RA Nukhopadhyay B., Stoddard S.F., Wolfe R.S.;
 RT "Purification, regulation, and molecular and biochemical
 RT characterization of pyruvate carboxylase from Methanobacterium
 RT thermoautotrophicum strain deltaH."
 RL J. Biol. Chem. 273:5155-5166(1998).
 CC -1 FUNCTION: PYRUVATE CARBOXYLASE CATALYZES A 2-STEP REACTION,
 CC INVOLVING THE ATP-DEPENDENT CARBOXYLATION OF THE COVALENTLY
 CC ATTACHED BIOTIN IN THE FIRST STEP AND THE TRANSFER OF THE CARBOXYL
 CC GROUP TO PYRUVATE IN THE SECOND. THE MAXIMUM ACTIVITY IS AT PH 8
 CC AND 60 DEGREES CELSIUS.
 CC -1 CATALYTIC ACTIVITY: ATP + pyruvate + HCO(3)(-) - ADP + phosphate +
 CC oxaloacetate.
 CC -1 COFACTOR: ATP, MAGNESIUM (OR MANGANESE OR COBALT), PYRUVATE AND
 CC BICARBONATE.
 CC -1- ENZYME REGULATION: INHIBITED BY ADP AND ALPHA-KETOGLUTARATE.
 CC -1 SUBUNIT: HETEROOCTAMER OF FOUR A AND FOUR B SUBUNITS.
 CC -1 SIMILARITY: WITH OTHER BIOTIN CARBOXYLASES, LIPOMIDE TRANSFERASES
 CC AND CARBAMYL PHOSPHATE SYNTHETASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AE000881; BAB85596.1; -
 DR EMBL: AF039105; AAC12719.1; -
 DR HSSP: P02905; IBDO.
 DR InterPro: IPR001882; Biotin_attach.
 DR InterPro: IPR000089; Biotin_lipoyl.
 DR InterPro: IPR000891; HMG-Like.
 DR InterPro: IPR003379; PYC_OADA.
 DR Pfam: PF00364; biotin_lipoyl.1.
 DR Pfam: PF02436; PYC_OADA.1.
 DR Pfam: PF02436; HMG-Like.1.
 DR TIGRPFAMs: TIGR01108; oada.1.
 DR PROSITE: PS00188; BIOTIN; FALSE_NEG.
 KW Ligase; Multifunctional enzyme; Glucocoenesals; Magnesium; Pyruvate;
 KW Biotin; Complete proteome.

FT BINDING 54 69 PYRUVATE.
FT BINDING 534 534 BIOTIN.
SQ SEQUENCE 568 AA; 63955 MW; D328715AB0328DBB CRC64;

Query Match 40.9%; Score 140; DB 1; Length 568;
Best Local Similarity 43.5%; Pred. No. 8.1e-07;
Matches 30; Conservative 11; Mismatches 28; Indels 0; Gaps 0;

OY 1 EGEIPIAGTIVSKIVLVEGDTVKGQTVLVLEAMKMETEINAPDGVKVEKVEDAV 60
DB 499 EGAIVSTQGMVYKLVKSVSGDQVNMGDVAVYEAAMKEMNDITPHGVYKITYTAEGEKV 558
OY 61 QGGGGLIKI 69
DB 559 EGDIIIMVI 567

RESULT 11
BCCP_PORPU STANDARD; PRT; 157 AA.

AC P51283;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Biotin carboxyl carrier protein of acetyl-CoA carboxylase (BCCP).
GN ACB.
OS Porphyra purpurea.
OC Chloroplast.
OC Eukaryota; Rhodophyta; Bangiophyceae; Bangiales; Bangiaceae; Porphyra.

OX NCBI_TaxID=2787;
RN [1]
RP SEQUENCE FROM N.A.

RA Reith M.E., Munroland J.;
RC STRAIN=Avonport;
RT "Complete nucleotide sequence of the Porphyra purpurea chloroplast genome."
RL Plant Mol. Biol. Rep. 13:333-335(1995).

CC -1- FUNCTION: THIS PROTEIN IS A COMPONENT OF THE ACETYL COENZYME A
CARBOXYLASE COMPLEX: FIRST, BIOTIN CARBOXYLASE CATALYZES THE
CARBOXYLATION OF THE CARRIER PROTEIN AND THEN THE TRANSCARBOXYLASE
TRANSFERS THE CARBOXYL GROUP TO FORM MALONYL-COA.
CC -1- PATHWAY: long-chain fatty acid biosynthesis; first step.

CC -1- SUBCELLULAR LOCATION: Chloroplast.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

CC EMBL: U38804; AAC08169.1; -

DR HSSP: P02905; IBD0.

DR InterPro: IPR001249; ACCOA_biotinc.

DR InterPro: IPR001882; Biotin_attach.

DR InterPro: IPR000089; Biotin_lipoyl.

DR Pfam: PF00364; Biotin_lipoyl; 1.

DR PRINTS: PR01071; ACOABIOTINC.

DR TIGRfams: TIGR00531; BCCP.1.

DR PROSITE: PS00188; BIOTIN; 1.

DR Fatty acid biosynthesis; Biotin; Chloroplast.

FT BINDING 122 122 BIOTIN (BY SIMILARITY).
SQ SEQUENCE 157 AA; 17519 MW; SCF079B2410E777 CRC64;

Query Match 40.5%; Score 138.5; DB 1; Length 157;
Best Local Similarity 42.5%; Pred. No. 3.1e-07;
Matches 31; Conservative 12; Mismatches 23; Indels 7; Gaps 1;

OY 4 IPADLAGTV-----SKLVKSGTVKAGQTVLVLEAMKMETEINAPDGVKVEKVEK 56
DB 83 IVPMVNGTFYHSPAPGEKIFQVGDIVKCNQTVCIIEAMKMLEALEEGIIIEILVKN 142

OY 57 RDAVGGGGLIKI 69
DB 143 GDIVDCGQALMKV 155

RESULT 12
PYC2_YEAST STANDARD; PRT; 1180 AA.

ID PYC2_YEAST
AC P32327;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Pyruvate carboxylase 2 (EC 6.4.1.1) (Pyruvic carboxylase 2) (PCB 2).
GN PYC2 OR YBR218C OR YBR1507.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.

OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A. AND CHARACTERIZATION.

RA MEDLINE=9201767; PubMed=1921979;
RA Stucka R., Deguin S., Salmon J.-M., Gancedo C.;

RT "DNA sequences in chromosomes II and VII code for pyruvate
carboxylase isoenzymes in Saccharomyces cerevisiae: analysis of
pyruvate carboxylase-deficient strains.";

RT Mol. Gen. Genet. 229:307-315(1991).

RL [2]
RN [1]
RP SEQUENCE FROM N.A.

RA STRAIN=S288C;
RA Dubois E., el Bakoury M., Glansdorff N., Messenguy F., Pierard A.,

RA Scherrens B., Vierendeels F.;

RT Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.

RN [3]
RP SEQUENCE FROM N.A.

RA MEDLINE=96128067; PubMed=8554526;
RA Val D.L., Chapman-Smith A., Walker M.E., Cronan J.E. Jr.,

RA Wallace J.C.;

RT "Polymorphism of the yeast pyruvate carboxylase 2 gene and protein:
effects on protein biotinylation.";

RT Biochem. J. 312:817-825(1995).

CC -1- FUNCTION: PYRUVATE CARBOXYLASE CATALYZES A 2-STEP REACTION,
INVOLVING THE ATP-DEPENDENT CARBOXYLATION OF THE COVALENTLY
ATTACHED BIOTIN IN THE FIRST STEP AND THE TRANSFER OF THE
CARBOXYL GROUP TO PYRUVATE IN THE SECOND.

CC -1- CATALYTIC ACTIVITY: ATP + pyruvate + HCO(3)(-) = ADP + phosphate +
oxaloacetate.

CC -1- COFACTOR: BIOTIN AND ZINC.

CC -1- PATHWAY: GLUCONEOGENESIS.

CC -1- SUBUNIT: HOMOTETRAMER.

CC -1- SUBCELLULAR LOCATION: Cytoplasmic.

CC -1- INDUCTION: BY GLUCOSE.

CC AND CARBAMYL PHOSPHATE SYNTHETASES.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

CC EMBL: X59890; CAA42544.1; -

DR EMBL: U36087; CAA85182.1; -

DR EMBL: U35647; AAC49147.1; -

DR PIR: S46094; S46094.

DR HSSP: P24182; IBCG.

DR SGD: S0000422; PYC2.
DR InterPro: IPR001882; Biotin_attach.
DR InterPro: IPR000089; Biotin_lipoyl.
DR InterPro: IPR000901; CPSase.
DR InterPro: IPR000891; HMG_L-like.
DR InterPro: IPR003379; PYC_OADA.

DR pfam: PF00289; CPSase_L_chain; 1.
 DR pfam: PF00364; biotin_lipoyl; 1.
 DR pfam: PF00682; HMG_L-like; 1.
 DR pfam: PF02436; PYC_ODA; 1.
 DR pfam: PF02785; biotin_carb_C; 1.
 DR pfam: PF02786; CPSase_L_D2; 1.
 DR TIGRfams; TIGR01235; pyruv_carbox; 1.
 DR PROSITE; PS00188; BIOTIN; 1.
 DR PROSITE; PS00866; CPSASE_1; 1.
 DR PROSITE; PS00867; CPSASE_2; 1.
 KW ligase; Multifunctional enzyme; Biotin; gluconeogenesis; ATP-binding;
 KW zinc; Multigene family.
 FT NP_BIND 183 188
 FT ACT_SITE 313 313
 FT BINDING 1136 1136
 FT SIMILAR 161 331
 FT SIMILAR 351 471
 FT SIMILAR 1087 1180
 FT CONFLICT 15 15
 FT CONFLICT 132 132
 FT CONFLICT 238 238
 FT CONFLICT 268 268
 FT CONFLICT 546 546
 FT CONFLICT 642 642
 FT CONFLICT 771 773
 FT CONFLICT 831 831
 FT CONFLICT 839 839
 FT CONFLICT 1001 1001
 FT CONFLICT 1155 1155
 FT CONFLICT 1178 1178
 FT CONFLICT 1180 1180
 SQ SEQUENCE 1180 AA; 130166 MW; AD60DA3A60F5E001 CRC64;
 Query Match Best Local Similarity 40.4%; Score 138; DB 1; Length 1180;
 Matches 25; Conservative 19; Mismatches 21; Indels 0; Gaps 0;
 QY 3 EIPAPLAGTYSKILVKGSDTVKAGQTVLVLEAMKMETEINAPTDGKVEKYLKERDVAOG 62
 DB 1103 QIGAPVAGYIEVKVHKSLVKKGESIAVLAMKMEMVYSSPADGVKNDFINDGESVDA 1162
 QY 63 GGGI 67
 DB 1163 SDLV 1167
 RESULT 13
 PYCL_YEAST STANDARD; PRT; 1178 AA.
 AC P11154;
 DT 01-JUL-1989 (Rel. 11, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Pyruvate carboxylase 1 (EC 6.4.1.1) (Pyruvic carboxylase 1) (PCB 1).
 GN PYCL OR PYV OR YG1062W.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RX MEDLINE=88298805; PubMed=3042770;
 RA Lim F., Morris C.P., Occhiodoro F., Wallace J.C.;
 RT "Sequence and domain structure of yeast pyruvate carboxylase.";
 RL J. Biol. Chem. 263:11493-11497(1988).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C;
 RX MEDLINE=97377993; PubMed=9234674;
 RA Feuerhahn M., de Montigny J., Potier S., Souciet J.-L.;
 RT "The characterization of two new clusters of duplicated genes
 suggests a 'Lego' organization of the yeast Saccharomyces cerevisiae

RT chromosomes.";
 RL Yeast 13:861-869(1997).
 RN [3]
 RP SEQUENCE OF 1003-1178 FROM N.A.
 RX MEDLINE=87241529; PubMed=3036126;
 RA Morris C.P., Lim F., Wallace J.C.;
 RT "Yeast pyruvate carboxylase: gene isolation.";
 RL Biochem. Biophys. Res. Commun. 145:390-396(1987).
 CC -1- FUNCTION: PYRUVATE CARBOXYLASE CATALYZES A 2-STEP REACTION,
 INVOLVING THE ATP-DEPENDENT CARBOXYLATION OF THE COVALENTLY
 ATTACHED BIOTIN IN THE FIRST STEP AND THE TRANSFER OF THE
 CARBOXYL GROUP TO PYRUVATE IN THE SECOND.
 CC -1- CATALYTIC ACTIVITY: ATP + pyruvate + HCO(3)(-) - ADP + phosphate +
 oxaloacetate.
 CC -1- COFACTOR: BIOTIN AND ZINC.
 CC -1- PATHWAY: GLUCONEOGENESIS.
 CC -1- SUBUNIT: HOMOTETRAMER.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- SIMILARITY: WITH OTHER BIOTIN CARBOXYLASES, LIPOAMIDE TRANSFERASES
 AND CARBAMYL PHOSPHATE SYNTHETASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL; J03889; AAA3484.1; -
 DR EMBL; Z72584; CAA96765.1; -
 DR PIR; A29233; OYBP.
 DR HSSP; P24182; 1BNC.
 DR S0003030; PYCL.
 DR InterPro; IPR001882; Biotin_attach.
 DR InterPro; IPR000089; Biotin_lipoyl.
 DR InterPro; IPR000901; CPSase.
 DR InterPro; IPR000891; HMG_L-like.
 DR InterPro; IPR003379; PYC_ODA.
 DR pfam: PF00289; CPSase_L_chain; 1.
 DR pfam: PF00364; biotin_lipoyl; 1.
 DR pfam: PF00682; HMG_L-like; 1.
 DR pfam: PF02436; PYC_ODA; 1.
 DR pfam: PF02785; biotin_carb_C; 1.
 DR pfam: PF02786; CPSase_L_D2; 1.
 DR TIGRfams; TIGR01235; pyruv_carbox; 1.
 DR PROSITE; PS00188; BIOTIN; 1.
 DR PROSITE; PS00866; CPSASE_1; 1.
 DR PROSITE; PS00867; CPSASE_2; 1.
 KW ligase; Multifunctional enzyme; Biotin; gluconeogenesis; ATP-binding;
 KW zinc; Multigene family.
 FT NP_BIND 182 187
 FT ACT_SITE 312 312
 FT BINDING 1135 1135
 FT SIMILAR 160 330
 FT SIMILAR 350 470
 FT SIMILAR 1086 1178
 FT CONFLICT 462 462
 FT CONFLICT 493 493
 FT CONFLICT 595 595
 FT CONFLICT 619 619
 FT CONFLICT 664 664
 FT CONFLICT 772 772
 FT CONFLICT 879 879
 FT CONFLICT 909 909
 SQ SEQUENCE 1178 AA; 130099 MW; BC7110A8AFB23E04 CRC64;
 Query Match Best Local Similarity 38.2%; Score 133; DB 1; Length 1178;
 Matches 27; Conservative 15; Mismatches 22; Indels 0; Gaps 0;
 QY 4 IPAPLAGTYSKILVKGSDTVKAGQTVLVLEAMKMETEINAPTDGKVEKYLKERDVAOG 63

Db 1103 IGAPMAGVVEVKHSGSLIKGQPPAVLSAKRMEMITISSPSDQGVKEVFSVDGENVDS 1162
 Qy 64 OGLI 67
 Db 1163 DLTV 1166

RESULT 14
 PYC_PICPA STANDARD: PRT: 1189 AA.
 AC P78992:
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUN-1999 (Rel. 38, Last annotation update)
 DE Pyruvate carboxylase (EC 6.4.1.1) (Pyruvic carboxylase) (PCB).
 GN PYC1.
 OS Pichia pastoris (Yeast).
 CC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 CC Saccharomycetales; Saccharomycetaceae; Pichia.
 OX NCBI_TaxID=4922;
 RN [1]
 RP MEDLINE=98301182; PubMed=9639311;
 RA Menendez J., Delgado J., Gancedo C.;
 RT "Isolation of the Pichia pastoris PYC1 gene encoding pyruvate
 carboxylase and identification of a suppressor of the pyc
 phenotype.";
 RT Yeast 14:647-654(1998).
 CC -1- FUNCTION: PYRUVATE CARBOXYLASE CATALYZES A 2-STEP REACTION,
 INVOLVING THE ATP-DEPENDENT CARBOXYLATION OF THE COVALENTLY
 ATTACHED BIOTIN IN THE FIRST STEP AND THE TRANSFER OF THE
 CARBOXYL GROUP TO PYRUVATE IN THE SECOND (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: ATP + Pyruvate + HCO(3)(-) - ADP + phosphate +
 oxaloacetate.
 CC -1- COFACTOR: BIOTIN AND ZINC.
 CC -1- PATHWAY: GLUCONEOGENESIS.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- SIMILARITY: WITH OTHER BIOTIN CARBOXYLASES, LIPOAMIDE TRANSFERASES
 AND CARBAMOYL PHOSPHATE SYNTHETASES.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: Y1106; CAA71993.1; -.
 DR HSSP: P24182; 1DVI.
 DR InterPro: IPR001882; Biotin_attach.
 DR InterPro: IPR000089; Biotin_lipoyl.
 DR InterPro: IPR000901; CPSase.
 DR InterPro: IPR000891; HMGL-like.
 DR InterPro: IPR003373; PYC_OXA.
 DR Pfam: PF00289; CPSase_L-chain; 1.
 DR Pfam: PF00364; Biotin_lipoyl; 1.
 DR Pfam: PF00682; HMGL-like; 1.
 DR Pfam: PF02436; PYC_OXA; 1.
 DR Pfam: PF02785; Biotin_carb_C; 1.
 DR Pfam: PF02786; CPSase_L_D2; 1.
 DR TIGRfams: TIGR01235; pyruv_carbox; 1.
 DR PROSITE: PS00188; BIOTIN; 1.
 DR PROSITE: PS00866; CPSASE_1; 1.
 DR PROSITE: PS00867; CPSASE_2; 1.
 KW Ligase; Multifunctional enzyme; Biotin; Gluconeogenesis; ATP-binding;
 KW Zinc.
 FT ACT_SITE 185 190 ATP (POTENTIAL).
 FT BINDING 315 315 BY SIMILARITY.
 FT BINDING 1140 1140 BIOTIN (BY SIMILARITY).
 SO SEQUENCE 1189 AA; 131400 MW; 8B6E856079657914 CRC64;

Query Match 38.6%; Score 132; DB 1; Length 1189;
 Best Local Similarity 46.6%; Pred. No. 1.1e-05;
 Matches 27; Conservative 11; Mismatches 20; Indels 0; Gaps 0;

Db 1107 EIGAPMAGVVEVREHNGEVRKGPDIIVLSAKRMKEMVSSVPAIGRIQIVKENDSV 1164
 Qy 3 EIRPILACTVSKIIIVKEDDFVACGTVLVLEAMKETFINAPTCKVKKVLYKENDAV 60
 Db 1107 EIGAPMAGVVEVREHNGEVRKGPDIIVLSAKRMKEMVSSVPAIGRIQIVKENDSV 1164

RESULT 15
 BCCA_MYCLE STANDARD: PRT: 598 AA.
 AC P46392:
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Acetyl-/propionyl-coenzyme A carboxylase alpha chain [includes: Biotin
 carboxylase (EC 6.3.4.14); Biotin carboxyl carrier protein (BCCP)].
 GN BCCA OR M10726 OR B1308_C1_129.
 OS Mycobacterium leprae.
 CC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
 CC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1769;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=94222829; PubMed=7909542;
 RA Norman E., de Smet K.A.L., Stoker N.G., Ratledge C., Wheeler P.R.,
 RA Dale J.W.;
 RT "Lipid synthesis in mycobacteria: characterization of the biotin
 carboxyl carrier protein genes from Mycobacterium leprae and M.
 tuberculosis.";
 RT J. Bacteriol. 176:2525-2531(1994).
 CC (2)
 RP SEQUENCE FROM N.A.
 RA Smith D.R., Robison K.;
 RL Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=TN;
 RX MEDLINE=21128732; PubMed=11234002;
 RA Cole S.T., Eigmeier K., Parkhill J., James K.D., Thomson N.R.,
 RA Wheeler P.R., Honore N., Garner T., Churcher C., Harris D.,
 RA Mungall K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Daniels R.M., Devlin K., Duthoy S., Felwell T., Fraser A., Hamlin N.,
 RA Houtby L., Hornsby T., Jagels K., Lacroix C., Mclean J., Moule S.,
 RA Murphy L., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,
 RA Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,
 RA Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,
 RA Batteil B.G.;
 RT "Massive gene decay in the leprosy bacillus.";
 RT Nature 409:1007-1011(2001).
 RL Nature 409:1007-1011(2001).
 CC -1- FUNCTION: THIS PROTEIN CARRIES TWO FUNCTIONS: BIOTIN CARBOXYL
 CARRIER PROTEIN AND BIOTIN CARBOXYLTRANSFERASE.
 CC -1- CATALYTIC ACTIVITY: ATP + biotin-carboxyl-carrier protein + CO(2)
 -> ADP + phosphate + carboxybiotin-carboxyl-carrier protein.
 CC -1- COFACTOR: BIOTIN.
 CC -1- PATHWAY: Long-chain fatty acid biosynthesis; first step.
 CC -1- SUBUNIT: MULTIMER COMPRISED OF 2 DIFFERENT SUBUNTS; THE LARGER
 ONE (63/64 kDa) HAS BIOTIN CARBOXYLASE AND BIOTIN CARRIER
 FUNCTIONS, WHILE THE SMALLER SUBUNIT POSSESSES CARBOXYLTRANSFERASE
 AND SUBSTRATE BINDING ACTIVITY.
 CC -1- SIMILARITY: TO OTHER BIOTIN-DEPENDENT ENZYMES AND CARBAMOYL-
 PHOSPHATE SYNTHETASES.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X63470; CAA45070.1; -.

DR EMBL: U00012: AA85920.1; -
 DR EMBL: AL583919; CAC30235.1; -
 DR HSSP: P24182; 1BNC.
 DR Leprosoma; ML0726; -
 DR InterPro: IPR001882; Biotin_attach.
 DR InterPro: IPR000089; Biotin_1lpoyl.
 DR InterPro: IPR000901; CPSase.
 DR Pfam: PF00289; CPSase_L_chain; 1.
 DR Pfam: PF00364; Biotin_1lpoyl; 1.
 DR Pfam: PF02785; Biotin_carb_C; 1.
 DR Pfam: PF02786; CPSase_L_D2; 1.
 DR PRINTS: PR00098; CPSASE.
 DR PROSITE: PS00188; BIOTIN; 1.
 DR PROSITE: PS00867; CPSASE_1; 1.
 DR PROSITE: PS00867; CPSASE_2; 1.
 DR Fatty acid biosynthesis; Ligase; Biotin; Multifunctional enzyme;
 KW ATP-binding; Complete proteome.
 FT DOMAIN 11 441 BIOTIN CARBOXYLASE
 FT NP_BIND 532 598 BIOTIN CARBOXYL CARRIER PROTEIN.
 FT ACT_SITE 169 174 ATP (BY SIMILARITY).
 FT BINDING 299 299 BY SIMILARITY.
 FT BINDING 564 564 BIOTIN (BY SIMILARITY).
 FT CONFLICT 30 30 D -> H (IN REF. 1).
 SQ SEQUENCE 598 AA: 63863 MW: 5F2E291D7C54515D CRC64;

Query Match 38.0%; Score 130; DB 1; Length 598;
 Best Local Similarity 43.9%; Pred. No. 8.6e-06;
 Matches 29; Conservative 10; Mismatches 27; Indels 0; Gaps 0;

OY 4 IPAPLAGVSKILVKEGDTYKAGQVLYLEAMKETEINAPTDGKYEKLVKERDAV 63
 DB 532 VTAPOGTVKVAAGQVTMTGDLVYLEAMKEMENPAHAKDGLITGLAVEGTAITGG 591

OY 64 OGLIRI 69
 DB 592 TVLAEL 597

RESULT 16
 ID: MCCA_SOYBN STANDARD; PRT: 731 AA.
 AC 042777; 042778;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Methylcrotonyl-CoA carboxylase alpha chain, mitochondrial precursor
 DE (EC 6.4.1.4) (3-Methylcrotonyl-CoA carboxylase 1) (MCCase alpha
 subunit) (3-methylcrotonyl-CoA:carbon dioxide ligase alpha subunit).
 GN MCCA.
 OS Glycine max (Soybean).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
 OX NCBI_TaxID=3847;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF N-TERMINUS.
 RC STRAIN=cv. Corsoy 79; TISSUE=Colyledon;
 RX MEDLINE=94286521; PubMed=8016064;
 RA Song J., Wirtle E.S., Nikolau B.J.;
 RT "Molecular cloning and characterization of the cDNA coding for the
 RT biotin-containing subunit of 3-methylcrotonyl-CoA carboxylase:
 RT Identification of the biotin carboxylase and biotin-carrier
 RT domains.";
 RT Proc. Natl. Acad. Sci. U.S.A. 91:5779-5783(1994).
 CC -1- CATALYTIC ACTIVITY: ATP + 3-methylcrotonyl-CoA + HCO(3)(-) = ADP +
 CC phosphate + 3-methylglutaconyl-CoA.
 CC -1- CORFACTOR: Biotin.
 CC -1- PATHWAY: Leucine catabolism.
 CC -1- SUBUNIT: Probably a heterodimer composed of biotin-containing
 CC alpha subunits and beta subunits (By similarity).
 CC -1- CELLULAR LOCATION: Mitochondrial matrix.
 CC -1- TISSUE SPECIFICITY: In leaves, cotyledons and stems.
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>
 CC or send an email to license@sdb.ch).

DR EMBL: U08469; AA53140.1; -
 DR EMBL: U08469; AA53141.1; -
 DR HSSP: P24182; 1BNC.
 DR InterPro: IPR001882; Biotin_attach.
 DR InterPro: IPR000089; Biotin_1lpoyl.
 DR InterPro: IPR000901; CPSase.
 DR Pfam: PF02785; Biotin_carb_C; 1.
 DR Pfam: PF00364; Biotin_1lpoyl; 1.
 DR Pfam: PF00289; CPSase_L_chain; 1.
 DR Pfam: PF02786; CPSase_L_D2; 1.
 DR PROSITE: PS00188; BIOTIN; 1.
 DR PROSITE: PS00867; CPSASE_1; 1.
 DR PROSITE: PS00867; CPSASE_2; 1.
 KW Mitochondrion; Ligase; Biotin; ATP-binding; Transf. peptide.
 FT TRANSIT 1 22 MITOCHONDRION
 FT CHAIN 23 731 METHYLCROTONYL-CoA CARBOXYLASE ALPHA
 FT NP_BIND 193 198 CHAIN.
 FT ACT_SITE 324 324 ATP (POTENTIAL).
 FT BINDING 695 695 BIOTIN (BY SIMILARITY).
 FT CONFLICT 68 68 R -> K (IN REF. 1; AA53141).
 FT CONFLICT 75 75 T -> S (IN REF. 1; AA53141).
 FT CONFLICT 78 78 E -> K (IN REF. 1; AA53141).
 SQ SEQUENCE 731 AA: 80619 MW: C4D5A94F8123A9B4 CRC64;

Query Match 38.0%; Score 130; DB 1; Length 731;
 Best Local Similarity 42.0%; Pred. No. 1.1e-05;
 Matches 29; Conservative 12; Mismatches 26; Indels 0; Gaps 0;

OY 1 ESEIPAPLAGVSKILVKEGDTYKAGQVLYLEAMKETEINAPTDGKYEKLVKERDAV 60
 DB 660 OGTVAPMAGLVKLVKENTRVEEGPVLYLEAMKEMENPAHAKDGLITGLAVEGTAITGG 719

OY 61 OGGGGLIRI 69
 DB 720 SDGSVLEFV 728

RESULT 17
 ID: PYC_MOUSE STANDARD; PRT: 1178 AA.
 AC 005920;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Pyruvate carboxylase, mitochondrial precursor (EC 6.4.1.1) (Pyruvic
 carboxylase) (PCB).
 GN PC OR PCX.
 OS Mus musculus (Mouse).
 OC Eukaryota; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Euteria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=adipocyte;
 RX MEDLINE=93189578; PubMed=8446588;
 RA Zhang J., Xia W.L., Brew K., Ahmad F.;
 RT "Adipose pyruvate carboxylase: amino acid sequence and domain
 RT structure deduced from cDNA sequencing.";
 RT Proc. Natl. Acad. Sci. U.S.A. 90:1766-1770(1993).
 CC -1- FUNCTION: PYRUVATE CARBOXYLASE CATALYZES A 2-STEP REACTION,
 CC INVOLVING THE ATP-DEPENDENT CARBOXYLATION OF THE COVALENTLY
 CC ATTACHED BIOTIN IN THE FIRST STEP AND THE TRANSFER OF THE
 CC CARBOXYL GROUP TO PYRUVATE IN THE SECOND. CATALYZES IN A TISSUE
 CC SPECIFIC MANNER, THE INITIAL REACTIONS OF GLUCOSE (LIVER, KIDNEY)

CC AND LIPID (ADIPOSE TISSUE, LIVER, BRAIN) SYNTHESIS FROM PYRUVATE.
CC -1- CATALYTIC ACTIVITY: ATP + pyruvate + HCO₃(-)- ADP + phosphate +
CC oxalacetate.
CC -1- COFACTOR: BIOTIN AND MANGANESE.
CC -1- PATHWAY: GLUCONEOGENESIS AND LIPOGENESIS.
CC -1- SUBUNIT: HOMOTETRAMER.
CC -1- SUBCELLULAR LOCATION: Mitochondrial matrix.
CC -1- TISSUE SPECIFICITY: LIVER, KIDNEY, ADIPOSE TISSUE, LIVER
CC AND BRAIN.
CC -1- SIMILARITY: WITH OTHER BIOTIN CARBOXYLASES, LIPAMIDE TRANSFERASES
CC AND CARBAMYL PHOSPHATE SYNTHETASES.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL: L09192; AAA39737.1; -.
CC PIR: A47255; A47255.
CC HSSP: P24182; INNC.
CC SWISS-2DPAGE: 005920; MOUSE.
CC MGD: MGI:97520; PCX.
CC InterPro: IPR001882; Biotin_attach.
CC InterPro: IPR000089; Biotin_lipoyl.
CC InterPro: IPR000901; CPSase.
CC InterPro: IPR000891; HMGL-like.
CC InterPro: IPR003379; PYC_OADA.
CC Pfam: PF00289; CPSase_L_chain; 1.
CC Pfam: PF00364; Biotin_lipoyl; 1.
CC Pfam: PF00682; HMGL-like; 1.
CC Pfam: PF02436; PYC_OADA; 1.
CC Pfam: PF02785; Biotin_carb_C; 1.
CC Pfam: PF02786; CPSase_L_D2; 1.
CC TIGRfams: TIGR01235; pyruv_carbox; 1.
CC PROSITE: PS00188; BIOTIN; 1.
CC
CC KW Ligase; Multifunctional enzyme; Biotin; Manganese; gluconeogenesis;
CC ATP-binding; Mitochondrion; Lipid synthesis; Transit peptide.
CC FT TRANSIT 1 20 MITOCHONDRION (POTENTIAL).
CC FT CHAIN 21 1178 PYRUVATE CARBOXYLASE.
CC FT DOMAIN 21 549 BIOTIN CARBOXYLASE (BY SIMILARITY).
CC FT DOMAIN 550 1000 CARBOXYLTRANSFERASE (BY SIMILARITY).
CC FT DOMAIN 1096 1178 BIOTIN CARBOXYL CARRIER PROTEIN
CC (BY SIMILARITY).
CC NP_BIND 198 203 ATP (BY SIMILARITY).
CC ACT_SITE 328 328 BIOTIN (BY SIMILARITY).
CC BINDING 1144 1144 BIOTIN (BY SIMILARITY).
CC SEQUENCE 1178 AA; 129684 MW; 14CEA0F9DA8B8127 CRC64;
SQ
Query Match 37.7%; Score 129; DB 1; Length 1178;
Best Local Similarity 40.6%; Pred. No. 2.1e-05;
Matches 28; Conservative 14; Mismatches 27; Indels 0; Gaps 0;
QY 1 EGEIPAPLAGVSKILVKEGDTVAKAGQTVLVLEAMKMETEINAPDGVVEVLYKERAV 60
Db 1109 KGQIGAPMPGVVIDIKVAAGDKVAKGQPLCVLSAMKMETVVTSPMEGTIRRVHTKDWTL 1168
QY 61 QGGGGLIKI 69
Db 1169 EGDLLIEI 1177
RESULT 18
BCCP_ANASP
ID BCCP_ANASP STANDARD; PRT; 181 AA.
AC Q06881;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Biotin carboxyl carrier protein of acetyl-CoA carboxylase (BCCP).
GN ACCB OR AL15057.

OS Anabaena sp. (strain PCC 7120).
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX NCBI_Taxid=103690;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-9.
RX MEDLINE=93352435; PubMed=8102363.
RA Gornicki P., Scappino L.A., Haselkorn R.;
RT "Genes for two subunits of acetyl coenzyme A carboxylase of Anabaena
sp. strain PCC 7120: biotin carboxylase and biotin carboxyl carrier
protein".
RL J. Bacteriol. 175:5268-5272(1993).
RL [2]
RN SEQUENCE FROM N.A.
RX MEDLINE=21595285; PubMed=11759840;
RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
RA Matsubara A., Iriguchi M., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA Nakazaki N., Shimo S., Sugimoto M., Takazawa M., Yamada M.,
RA Yasuda M., Tabata S.;
RT "Complete genomic sequence of the filamentous nitrogen-fixing
cyanobacterium Anabaena sp. strain PCC 7120".
RL DNA Res. 8:205-213(2001).
CC -1- FUNCTION: THIS PROTEIN IS A COMPONENT OF THE ACETYL COENZYME A
CC CARBOXYLASE COMPLEX, FIRST, BIOTIN CARBOXYLASE CATALYZES THE
CC TRANSFER OF THE CARBOXYL GROUP TO FORM MALONYL-COA.
CC -1- PATHWAY: Long-chain fatty acid biosynthesis; first step.
CC -1- SUBUNIT: HOMODIMER.
CC
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: L14863; AAA74628.1; -.
CC EMBL: AP003598; BAB76756.1; -.
CC HSSP: P02905; IBD0.
CC InterPro: IPR001249; ACCOA_biotinCC.
CC InterPro: IPR001882; Biotin_attach.
CC InterPro: IPR000089; Biotin_lipoyl.
CC Pfam: PF00364; Biotin_lipoyl; 1.
CC PRINTS: PR01071; AC0AB10TNC.
CC TIGRfams: TIGR00531; BCCP; 1.
CC PROSITE: PS00188; BIOTIN; 1.
CC Fatty acid biosynthesis; Biotin; Complete proteome.
CC INT_MET 0
CC BINDING 145 145 BIOTIN (BY SIMILARITY).
CC SEQUENCE 181 AA; 19049 MW; EBEC7B16C8D225F CRC64;
SQ
Query Match 37.0%; Score 126.5; DB 1; Length 181;
Best Local Similarity 35.1%; Pred. No. 5.8e-06;
Matches 26; Conservative 19; Mismatches 22; Indels 7; Gaps 1;
QY 3 EIPAPLAGVSK-----ILVKEGDTVAKAGQTVLVLEAMKMETEINAPDGVVEVLYK 55
Db 105 EVASPMVGTFRAPAPGPAVEVGRIRIQQGQVCIIEAMKIMNEIADVSGQVTEILVQ 164
QY 56 ERDAVGGGGLIKI 69
Db 165 NGEPVEYNQPLMRI 178
RESULT 19
BCCA_MYCTU
ID BCCA_MYCTU STANDARD; PRT; 654 AA.
AC P46401;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Acetyl-/propionyl-coenzyme A carboxylase alpha chain [includes: Biotin

DE carboxylase (EC 6.3.4.14); Biotin carboxyl carrier protein (BCCP)].
 GN ACBA1 OR BECA OR RV2501C OR MT2576 OR MTCY07A7.07C.
 OS Mycobacterium tuberculosis.
 CC Bacteria: Actinobacteria; Actinobacteria (class): Actinobacteridae;
 CC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 CC NCBI_TaxID=1773;
 RN [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=50410;
 RC MEDLINE=94222829; PubMed=7909542;
 RA Norman E., de Smet K.A.L., Stoker N.G., Ratledge C., Wheeler P.R.,
 RA Dale J.W.,
 RT "Lipid synthesis in mycobacteria: characterization of the biotin
 RT carboxyl carrier protein genes from Mycobacterium leprae and M.
 RT tuberculosis".
 RT J. Bacteriol. 176:2525-2531(1994).
 RN [2]
 RN SEQUENCE FROM N.A.
 RC STRAIN=H37RV;
 RC MEDLINE=98295987; PubMed=9634230;
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
 RA Gordon S.V., Eiglmeyer K., Gas S., Barry C.E. III, Tekala F.,
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holtroyd S.,
 RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
 RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
 RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.,
 RA "Deciphering the biology of Mycobacterium tuberculosis from the
 RT complete genome sequence".
 RT Nature 393:537-544(1998).
 RN [3]
 RN SEQUENCE FROM N.A.
 RC STRAIN=CDC 1551 / Oshkosh;
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
 RA Peterson J., DeBoy R., Dodson R., Gwin M.L., Haft D., Hickey E.,
 RA Kolonay J.F., Nelson M.C., Umayam L.A., Esmolaeva M.D., Salzberg S.L.,
 RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
 RA Bishtal W.,
 RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
 RT laboratory strains".
 RT Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 RL -1- FUNCTION: THIS PROTEIN CARRIES TWO FUNCTIONS: BIOTIN CARBOXYL
 CC CARRIER PROTEIN AND BIOTIN CARBOXYLTRANSFERASE.
 CC -1- CATALYTIC ACTIVITY: ATP + biotin-carboxyl-carrier protein + CO(2)
 CC = ADP + phosphate + carboxybiotin-carboxyl-carrier protein.
 CC -1- COFACTOR: BIOTIN.
 CC -1- PATHWAY: long-chain fatty acid biosynthesis; first step.
 CC -1- SUBUNIT: MULTIMER COMPRISED OF 2 DIFFERENT SUBUNITS, THE LARGER
 CC ONE (63/64 kDa) HAS BIOTIN CARBOXYLASE AND BIOTIN CARRIER
 CC FUNCTIONS, WHILE THE SMALLER SUBUNIT POSSESSES CARBOXYLTRANSFERASE
 CC AND SUBSTRATE BINDING ACTIVITY.
 CC -1- SIMILARITY: TO OTHER BIOTIN-DEPENDENT ENZYMES AND CARBAMOYL-
 CC PHOSPHATE SYNTHETASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: Z19549; CAA79609.1; -
 CC EMBL: Z19556; CAB08919.1; -
 CC EMBL: AE007094; AAK46880.1; -
 CC HSSP: P24182; 1BNC.
 CC TIGR: MT2576; -
 CC Tuberculist: RV2501C; -
 CC InterPro: IPR001882; Biotin_attach.
 CC InterPro: IPR000089; Biotin_lipoyl.
 CC InterPro: IPR000901; CPSSase.
 CC Pfam: PF00289; CPSSase_L_chain; 1.

DR Pfam: PF00364; biotin_lipoyl; 1.
 DR Pfam: PF02785; biotin_carb_C; 1.
 DR Pfam: PF02786; CPSSase_L_D2; 1.
 DR PROSITE: PS00188; BIOTIN; 1.
 DR PROSITE: PS00866; CPSSase_1; 1.
 DR PROSITE: PS00867; CPSSase_2; 1.
 KW Fatty acid biosynthesis; Ligase; Biotin; Multifunctional enzyme;
 KW ATP-binding; Complete proteome.
 FT DOMAIN 1 437 BIOTIN CARBOXYLASE.
 FT NP_BIND 588 654 BIOTIN CARBOXYL CARRIER PROTEIN.
 FT ACT_SITE 162 167 ATP (BY SIMILARITY).
 FT BINDING 294 294 BY SIMILARITY.
 FT ACTING 620 620 BIOTIN (BY SIMILARITY).
 SQ SEQUENCE 654 AA; 70592 MW; FAAD0A146432CABF CRC64;
 Query Match 36.4%; Score 124.5; DB 1; Length 654;
 Best Local Similarity 41.8%; Pred. No. 3.3e-05;
 Matches 28; Conservative 14; Mismatches 24; Indels 1; Gaps 1;
 QY 3 EIPAPLAGTVSKILVEGDTVKAGQTVLVLEANKMETEINAPDGVKEVYERDAVG 62
 DB 587 EVVSPMGSVIAYOVESGSGISAGDVYVVEANKMESHLEAPVSGRVQ-VLVSVSGQVAV 645
 QY 63 GGGIIRI 69
 DB 646 EGVLIARI 652
 RESULT 20
 BCCP_CHAIN
 ID BCCP_CHAIN STANDARD; PRT; 164 AA.
 AC 084125;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Biotin carboxyl carrier protein of acetyl-CoA carboxylase (BCCP).
 GN ACBA1 OR BECA OR RV2501C OR MT2576 OR MTCY07A7.07C.
 OS Mycobacterium tuberculosis.
 CC Bacteria: Actinobacteria; Actinobacteria (class): Actinobacteridae;
 CC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 CC NCBI_TaxID=1773;
 RN [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=H37RV;
 RC MEDLINE=98295987; PubMed=9634230;
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
 RA Gordon S.V., Eiglmeyer K., Gas S., Barry C.E. III, Tekala F.,
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holtroyd S.,
 RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
 RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
 RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.,
 RA "Deciphering the biology of Mycobacterium tuberculosis from the
 RT complete genome sequence".
 RT Nature 393:537-544(1998).
 RN [3]
 RN SEQUENCE FROM N.A.
 RC STRAIN=CDC 1551 / Oshkosh;
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
 RA Peterson J., DeBoy R., Dodson R., Gwin M.L., Haft D., Hickey E.,
 RA Kolonay J.F., Nelson M.C., Umayam L.A., Esmolaeva M.D., Salzberg S.L.,
 RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
 RA Bishtal W.,
 RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
 RT laboratory strains".
 RT Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 RL -1- FUNCTION: THIS PROTEIN CARRIES TWO FUNCTIONS: BIOTIN CARBOXYL
 CC CARRIER PROTEIN AND BIOTIN CARBOXYLTRANSFERASE.
 CC -1- CATALYTIC ACTIVITY: ATP + biotin-carboxyl-carrier protein + CO(2)
 CC = ADP + phosphate + carboxybiotin-carboxyl-carrier protein.
 CC -1- COFACTOR: BIOTIN.
 CC -1- PATHWAY: long-chain fatty acid biosynthesis; first step.
 CC -1- SUBUNIT: MULTIMER COMPRISED OF 2 DIFFERENT SUBUNITS, THE LARGER
 CC ONE (63/64 kDa) HAS BIOTIN CARBOXYLASE AND BIOTIN CARRIER
 CC FUNCTIONS, WHILE THE SMALLER SUBUNIT POSSESSES CARBOXYLTRANSFERASE
 CC AND SUBSTRATE BINDING ACTIVITY.
 CC -1- SIMILARITY: TO OTHER BIOTIN-DEPENDENT ENZYMES AND CARBAMOYL-
 CC PHOSPHATE SYNTHETASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AE001286; AAC67714.1; -
 CC HSSP: P02905; 3BDO.
 CC PHCI-2DPAGE; 084125; -
 CC InterPro: IPR001249; AcCoA_biotinCC.
 CC InterPro: IPR001882; Biotin_attach.
 CC InterPro: IPR000089; Biotin_lipoyl.
 CC Pfam: PF00364; biotin_lipoyl; 1.
 CC PRINTS: PR01071; ACOABIOITINCC.

DR TIGRFAMS: TIGR00531; BCCP; 1.
 DR PROSITE: PS00188; BIOTIN; 1.
 KW Fatty acid biosynthesis; Biotin; Complete proteome.
 FT BINDING 128 Biotin (BY SIMILARITY).
 SQ SEQUENCE 164 AA; 18198 MW; BDBACEBC2D394C CXC64;
 Query Match 35.7%; Score 122; DB 1; Length 164;
 Best Local Similarity 36.1%; Pred. NO. 1.5e-05;
 Matches 30; Conservative 11; Mismatches 28; Indels 14; Gaps 2;
 QY 1 EGEIPA-----PLAGTV-----SKLVKSGDTVAGGVVLEAMKMETEINAPTD 46
 Db 79 ESEAPAGDPLVSPVGVFGYSPSEAPAFIKPGDTVEDVYVCTEAMKVNVEYKAGMS 138
 QY 47 GRVENVYKEDAVOGGGLIKI 69
 Db 139 GRVEELITNGDPVOFGSKLFR 161
 RESULT 21
 PYC_HUMAN STANDARD; PRT; 1178 AA.
 AC P11496; Q16705;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Pyruvate carboxylase, mitochondrial precursor (EC 6.4.1.1) (Pyruvic
 DE carboxylase) (PCB).
 GN PC.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 ON NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver, and Kidney;
 RX MEDLINE=95002202; PubMed=7918683;
 RA Wexler I.D., Du Y., Lisgaris M.V., Mandal S.R., Freytag S.O.,
 RA Yang B.-S., Liu T.-C., Kwon M., Patel M.S., Kerr D.S.;
 RT "Primary amino acid sequence and structure of human pyruvate
 RT carboxylase.";
 RL Biochim. Biophys. Acta 1227:46-52(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney;
 RX MEDLINE=94324922; PubMed=8048912;
 RA Mackay N., Rigat B., Douglas C., Chen H.-S., Robinson B.H.;
 RT "cDNA cloning of human kidney pyruvate carboxylase.";
 RL Biochem. Biophys. Res. Commun. 202:1009-1014(1994).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver, and Kidney;
 RX Walker M.E., Jitrapakdee S., Val D.L., Wallace J.C.;
 RL Submitted (JUL-1995) to the EMBL/Genbank/DBJ databases.
 RN [4]
 RP SEQUENCE OF 1083-1178 FROM N.A.
 RX MEDLINE=87212051; PubMed=3555348;
 RA Lambornah A.-W., Quan F., Gravel R.A.;
 RT "Sequence homology around the biotin-binding site of human
 RT propionyl-CoA carboxylase and pyruvate carboxylase.";
 RL Arch. Biochem. Biophys. 254:631-636(1987).
 RN [5]
 RP SEQUENCE OF 1135-1178 FROM N.A.
 RX MEDLINE=85030380; PubMed=6548474;
 RA Freytag S.O., Collier R.J.;
 RT "Molecular cloning of a cDNA for human pyruvate carboxylase.
 RT Structural relationship to other biotin-containing carboxylases and
 RT regulation of mRNA content in differentiating preadipocytes.";
 RL J. Biol. Chem. 259:12831-12837(1984).
 RN [6]
 RP VARIANTS PC DEFICIENCY THR-610 AND ILE-743.
 RX MEDLINE=98254451; PubMed=9585612;
 RA Carbone M.A., Mackay N., Ling M., Cole D.E.C., Douglas C., Rigat B.,

RA Felgenbaum A., Clarke J.T.R., Haworth J.C., Greenberg C.R.,
 RA Seargent L., Robinson B.H.;
 RT "American pyruvate carboxylase deficiency is associated with two
 RT distinct missense mutations.";
 RL Am. J. Hum. Genet. 62:1312-1319(1998).
 CC -1- FUNCTION: PYRUVATE CARBOXYLASE CATALYZES A 2-STEP REACTION,
 CC INVOLVING THE ATP-DEPENDENT CARBOXYLATION OF THE COVALENTLY
 CC ATTACHED BIOTIN IN THE FIRST STEP AND THE TRANSFER OF THE
 CC CARBOXYL GROUP TO PYRUVATE IN THE SECOND. CATALYZES IN A TISSUE
 CC SPECIFIC MANNER, THE INITIAL REACTIONS OF GLUCOSE (LIVER, KIDNEY)
 CC AND LIPID (ADIPOSE TISSUE, LIVER, BRAIN) SYNTHESIS FROM PYRUVATE.
 CC -1- CATALYTIC ACTIVITY: ATP + pyruvate + HCO(3)(-) = ADP + phosphate +
 CC oxaloacetate.
 CC -1- COFACTOR: BIOTIN AND MANGANESE.
 CC -1- PATHWAY: GLUCONEOGENESIS AND LIPOGENESIS.
 CC -1- SUBUNIT: HOMOTETRAMER.
 CC -1- SUBCELLULAR LOCATION: Mitochondrial matrix.
 CC -1- DISEASE: DEFICIENCY IN PC CAUSES LACTIC ACIDOSIS, MENTAL
 CC RETARDATION AND DEATH, OCCURS IN THREE FORMS: TYPE A (MILD); TYPE
 CC B (SEVERE NEONATAL) AND A VERY MILD LACTICACIDEMIA.
 CC -1- SIMILARITY: WITH OTHER BIOTIN CARBOXYLASES, LIPOMIDE TRANSFERASES
 CC AND CARBAMYL PHOSPHATE SYNTHETASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: U04641; AAA95937.1; -
 DR EMBL: S72370; AAA31500.1; -
 DR EMBL: U30891; AAA82937.1; -
 DR EMBL: M26122; AAA36423.1; -
 DR EMBL: K02822; AAA60033.1; -
 DR PIR: B27883; B27883.
 DR PIR: S01469; S01469.
 DR HSSP: P24182; IBNC.
 DR Genew; HGNC:8636; PC.
 DR MIM: 266150; -
 DR Interpro: IPR001882; Biotin_attach.
 DR Interpro: IPR000089; Biotin_lipoyl.
 DR Interpro: IPR000901; CPSase.
 DR Interpro: IPR000891; HMG-Like.
 DR Interpro: IPR003379; Pyc-ODA.
 DR Pfam: PF00289; CPSase_L-chain; 1.
 DR Pfam: PF00364; biotin_lipoyl; 1.
 DR Pfam: PF00682; biotin_lipoyl; 1.
 DR Pfam: PF02436; Pyc-ODA; 1.
 DR Pfam: PF02785; Biotin_carb_C; 1.
 DR Pfam: PF02786; CPSase_L-D2; 1.
 DR TIGRFAMS: TIGR01235; pyruv-carbox; 1.
 DR PROSITE: PS00188; BIOTIN; 1.
 DR Lysase; Multifunctional enzyme; Biotin; Manganese; Gluconeogenesis;
 DR ATP-binding; Mitochondrion; Lipid synthesis; Transit peptide;
 DR Disease mutation.
 KW TRANSIT 1 20
 FT CHAIN 21 1178
 FT DOMAIN 21 549
 FT DOMAIN 550 1000
 FT DOMAIN 1096 1178
 FT NP_BIND 198 203
 FT ACT_SITE 328 328
 FT BINDING 1144 1144
 FT VARIANT 610 610
 FT VARIANT 743 743
 FT VARIANT 743 743
 FT CONFLICT 225 226
 FT CONFLICT 352 352
 FT CONFLICT 385 386
 M -> I (IN PC DEFICIENCY TYPE A).
 /FTID=VAR.008095.
 M -> T (IN PC DEFICIENCY TYPE A).
 /FTID=VAR.008096.
 LA -> WP (IN REF. 2).
 A -> S (IN REF. 3).
 RS -> PT (IN REF. 2).

FT -CONFLICT 486 487 EL -> DV (IN REF. 2).
 FT CONFLICT 638 638 P -> R (IN REF. 2).
 FT CONFLICT 729 729 E -> A (IN REF. 2).
 FT CONFLICT 774 775 DT -> AP (IN REF. 2).
 SQ SEQUENCE 1178 AA; 129633 MM; 381F527553AA0095 CMC64;
 Query Match 35.4%; Score 121; DB 1; Length 1178;
 Best Local Similarity 40.6%; Pred. No. 0.00014;
 Matches 28; Conservative 13; Mismatches 28; Indels 0; Gaps 0;
 QY 1 EGEIPPLACTVSKILYKEDGTAKAGTVLYLEAMKMETEINAPTDCKVEKVLKEDAY 60
 DB 1109 KGIQIGAMPKGVLDIKVYAKAKVAGKGPLCYLSAMKMEIVTSPMECTIRKHYTKDML 1168
 QY 61 QGGGGLIKI 69
 DB 1169 EGGDLILEI 1177
 RESULT 22
 ID PYC_RAT STANDARD; PRT; 1178 AA.
 AC P52873; 064555;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Pyruvate carboxylase, mitochondrial precursor (EC 6.4.1.1) (Pyruvic
 DE carboxylase) (PCB).
 GN PC.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RX MEDLINE=96096548; PubMed=8522203;
 RA Lehn D.A., Moran S.M., MacDonald M.J.;
 RT "The sequence of the rat pyruvate carboxylase-encoding cDNA";
 RL Gene 165:331-332(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Mistar; TISSUE=Liver;
 RX MEDLINE=96257760; PubMed=8687410;
 RA Jitrapakdee S., Booker G.M., Cassidy A.I., Wallace J.C.;
 RT "Cloning, sequencing and expression of rat liver pyruvate
 RT carboxylase";
 RL Biochem. J. 316:631-637(1996).
 CC -1- FUNCTION: PYRUVATE CARBOXYLASE CATALYZES A 2-STEP REACTION,
 CC INVOLVING THE ATP-DEPENDENT CARBOXYLATION OF THE COVALENTLY
 CC ATTACHED BIOTIN IN THE FIRST STEP AND THE TRANSFER OF THE
 CC CARBOXYL GROUP TO PYRUVATE IN THE SECOND. CATALYZES IN A TISSUE
 CC SPECIFIC MANNER, THE INITIAL REACTIONS OF GLUCOSE (LIVER, KIDNEY)
 CC AND LIPID (ADIPOSE TISSUE, LIVER, BRAIN) SYNTHESIS FROM PYRUVATE.
 CC -1- CATALYTIC ACTIVITY: ATP + Pyruvate + HCO(3)(-) = ADP + phosphate +
 CC oxaloacetate.
 CC -1- COFACTOR: BIOTIN AND MANGANESE (BY SIMILARITY).
 CC -1- PATHWAY: GLUCONEOGENESIS AND LIPOGENESIS.
 CC -1- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Mitochondrial matrix.
 CC -1- SIMILARITY: WITH OTHER BIOTIN CARBOXYLASES, LIPOAMIDE TRANSFERASES
 CC AND CARBAMYL PHOSPHATE SYNTHETASES.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC EMBL: U32314; AAA96256.1; -;
 DR EMBL: U36585; AAC52668.1; -;

DR HSSP; P24182; 1BNC.
 DR InterPro; IPR001882; Biotin_attach.
 DR InterPro; IPR000089; Biotin_11poyl.
 DR InterPro; IPR000901; CPSase.
 DR InterPro; IPR000891; HMGCL-1like.
 DR InterPro; IPR003379; PYC_OADA.
 DR Pfam; PF00289; CPSase_L_chain; 1.
 DR Pfam; PF00364; Biotin_11poyl; 1.
 DR Pfam; PF00682; HMGCL-1like; 1.
 DR Pfam; PF02436; PYC_OADA; 1.
 DR Pfam; PF02785; Biotin_carp_C; 1.
 DR Pfam; PF02786; CPSase_L_D2; 1.
 DR TIGRfams; TIGR01235; pyruv_carbox; 1.
 DR PROSITE; PS00188; Biotin; 1.
 KW Ligase; Multifunctional enzyme; Biotin; Manganese; Glucconeogenesis;
 KM ATP-Binding; Mitochondrion; Lipid synthesis; Transit peptide.
 FT TRANSIT 1 20 MITOCHONDRION (POTENTIAL).
 FT CHAIN 21 1178 PYRUVATE CARBOXYLASE.
 FT DOMAIN 21 549 BIOTIN CARBOXYLASE (BY SIMILARITY).
 FT DOMAIN 550 1000 CARBOXYLTRANSFERASE (BY SIMILARITY).
 FT DOMAIN 1096 1178 BIOTIN CARBOXYL CARRIER PROTEIN
 FT (BY SIMILARITY).
 FT NP_BIND 198 203 ATP (BY SIMILARITY).
 FT ACT_SITE 328 328 BIOTIN (BY SIMILARITY).
 FT BINDING 1144 1144 BIOTIN (BY SIMILARITY).
 FT CONFLICT 222 222 P -> S (IN REF. 2).
 FT CONFLICT 866 866 D -> I (IN REF. 2).
 FT CONFLICT 977 977 G -> R (IN REF. 2).
 SQ SEQUENCE 1178 AA; 129689 MM; 8E5FA19BC132A8DD CMC64;
 Query Match 34.8%; Score 119; DB 1; Length 1178;
 Best Local Similarity 37.7%; Pred. No. 0.00021;
 Matches 26; Conservative 15; Mismatches 28; Indels 0; Gaps 0;
 QY 1 EGEIPPLACTVSKILYKEDGTAKAGTVLYLEAMKMETEINAPTDCKVEKVLKEDAY 60
 DB 1109 KGIQIGAMPKGVLDIKVYAKAKVAGKGPLCYLSAMKMEIVTSPMECTIRKHYTKDML 1168
 QY 61 QGGGGLIKI 69
 DB 1169 EGGDLILEI 1177
 RESULT 23
 ID BTB7_MYCSM STANDARD; PRT; 70 AA.
 AC 09XCD6;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Biotinylated protein BTB7.3 homolog.
 OS Mycobacterium smegmatis.
 CC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
 CC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1772;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 700084 / mc(2)155;
 RX MEDLINE=99328972; PubMed=10400584;
 RA Fernandes N.D., Wu Q.-L., Kong D., Puyang X., Garg S., Husson R.N.;
 RT "A mycobacterial extracytoplasmic sigma factor involved in survival
 RT following heat shock and oxidative stress.";
 RL J. Bacteriol. 181:4266-4274(1999).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC EMBL: AF144091; AAD41812.1; -;

DR HSSP: P10802; 11YU. Biotin attach.
 DR InterPro: IPR001882; Biotin_attach.
 DR InterPro: IPR000089; Biotin_lipoyl.
 DR Pfam: PF00364; Biotin_lipoyl. 1.
 DR PROSITE: PS00188; Biotin; FALSE_NEG.
 KW Biotin.
 FT INIT MET 0 0 BY SIMILARITY.
 FT BINDING 36 36 BIOTIN (BY SIMILARITY).
 SQ SEQUENCE 70 AA; 7306 MW; 975C293B63C770C8 CRC64;
 Query Match 34.5%; Score 118; DB 1; Length 70;
 Best Local Similarity 39.3%; Pred. No. 1.6e-05;
 Matches 24; Conservative 14; Mismatches 23; Indels 0; Gaps 0;
 QY 3 EIPAPLAGYTKLVKGGTVKAGQVTVLEAKMKMETINAPTDGKVKYKVERDAVG 62
 DB 3 DVAEIVASVLEVYVHEGDIGEDTLVLESMKMEIPVLAEVAGTVTKVNAEGDVTA 62
 QY 63 G 63
 DB 63 G 63
 RESULT 24
 ID BCCP_CHLMU STANDARD; PRT; 163 AA.
 AC 09PKR5;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Biotin carboxyl carrier protein of acetyl-CoA carboxylase (BCCP).
 GN ACB OR TC0399.
 OS Chlamydia muridarum.
 OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
 OX NCBI_TaxId=83560;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MOPn / N199;
 RX MEDLINE=20150255; PubMed=10684935;
 RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heldelberg J.F.,
 RA White O., Hickey E.K., Peterson J., Ulteback T., Berry K., Bass S.,
 RA Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
 RA Gilm M., Nelson W., Deboy R., Kolonay J., McClarty G., Salzberg S.L.,
 RA Eisen J., Fraser C.M.;
 RT "Genome sequences of Chlamydia trachomatis Mopn and Chlamydia pneumoniae AR39.";
 RL Nucleic Acids Res. 28:1397-1406(2000).
 CC -1- FUNCTION: THIS PROTEIN IS A COMPONENT OF THE ACETYL COENZYME A CARBOXYLASE COMPLEX; FIRST, BIOTIN CARBOXYLASE CATALYZES THE CARBOXYLATION OF THE CARRIER PROTEIN AND THEN THE TRANSCARBOXYLASE TRANSFERS THE CARBOXYL GROUP TO FORM MALONYL-COA (By similarity).
 CC -1- PATHWAY: Long-chain fatty acid biosynthesis; first step.
 CC -1- SUBUNIT: HOMODIMER (By similarity).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC EMBL: AE002306; AAF39256.1;
 DR HSSP: P02905; IBDO.
 DR TIGR: TC0399;
 DR InterPro: IPR001249; AcCoA_biotinCC.
 DR InterPro: IPR001882; Biotin_attach.
 DR InterPro: IPR000089; Biotin_lipoyl.
 DR Pfam: PF00364; Biotin_lipoyl. 1.
 DR PRINTS: PR01071; ACOBBIOTINCC.
 DR TIGRFAMS: TIGR00531; BCCP. 1.
 DR PROSITE: PS00188; Biotin; 1.
 KW Fatty acid biosynthesis; Biotin; Complete proteome.

FT BINDING 127 127 BIOTIN (BY SIMILARITY).
 SQ SEQUENCE 163 AA; 18084 MW; 8536D58B5CE11D3B CRC64;
 Query Match 34.5%; Score 118; DB 1; Length 163;
 Best Local Similarity 40.3%; Pred. No. 3.7e-05;
 Matches 31; Conservative 10; Mismatches 28; Indels 8; Gaps 2;
 QY 1 EGE-IPAPLAGY-----SKLVKGGTVKAGQVTVLEAKMKMETINAPTDGKVKY 52
 DB 84 EGDIVSPVGVTEYVGPSPSPAFVPCDVISEDVVCIVEAKVNVKAGAGVVEV 143
 QY 53 LVKERDAVGGGGLIKI 69
 DB 144 LITNGDPVDFGSKLRI 160
 RESULT 25
 ID ODP2_PSEAE STANDARD; PRT; 547 AA.
 AC 059638;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Dihydrolipoamide acetyltransferase component of pyruvate dehydrogenase complex (EC 2.3.1.12) (E2).
 GN ACBF OR ACBB OR PA5016.
 OS Pseudomonas aeruginosa.
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 OC Pseudomonas.
 OX NCBI_TaxId=287;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=PAO;
 RX MEDLINE=9715227; PubMed=9171401;
 RA Rae J.L., Cutfield J.F., Lamont I.L.;
 RT "Sequences and expression of pyruvate dehydrogenase genes from Pseudomonas aeruginosa.";
 RL J. Bacteriol. 179:3361-3371(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 15692 / PAO1;
 RX MEDLINE=20437337; PubMed=10984043;
 RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,
 RA Hickey M.J., Brinkman F.S.L., Hutnagle W.O., Kowalik D.J., Lagrou M.,
 RA Gardner R.L., Goltry L., Tolentino E., Westbrook-Medmen S., Yuan Y.,
 RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Labig K., Lam R.M.,
 RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
 RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
 RT "Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic pathogen.";
 RL Nature 406:959-964(2000).
 CC -1- FUNCTION: The pyruvate dehydrogenase complex catalyzes the overall conversion of pyruvate to acetyl-CoA and CO(2). It contains multiple copies of three enzymatic components: pyruvate dehydrogenase (E1), dihydrolipoamide acetyltransferase (E2) and lipoamide dehydrogenase (E3).
 CC -1- CATALYTIC ACTIVITY: Acetyl-CoA + dihydrolipoamide = CoA + S-acetyldihydrolipoamide.
 CC -1- COFACTOR: THE E2 COMPONENT CONTAINS TWO COVALENTLY-BOUND LIPOYL COFACTORS (BY SIMILARITY).
 CC -1- SUBUNIT: FORMS A 24-POLYPEPTIDE STRUCTURAL CORE WITH OCTAHEDRAL SYMMETRY.
 CC SIMILARITY: BELONGS TO THE 2-OXOACID DEHYDROGENASE FAMILY.
 CC -1- SIMILARITY: CONTAINS 2 LIPOYL-BINDING DOMAINS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC

DR EMBL: U47920; AAC45354.1; -
DR EMBL: AE004914; AAC08401.1; -
DR HSSP: P10802; 1EAF.
DR InterPro: IPR001078; 20xocid.dh.
DR InterPro: IPR000089; Biotin_lipoyl.
DR InterPro: IPR004167; E3_binding.
DR InterPro: IPR003016; Lipoyl.
DR Pfam: PF00198; 2-oxoacid_dh; 1.
DR Pfam: PF00364; biotin_lipoyl; 2.
DR Pfam: PF02817; e3_binding; 1.
DR ProDom: PD001115; 20xocid_dh; 1.
DR PROSITE: PS00189; LIPOYL; 2.
KW Glycolysis; Transferase; Acyltransferase; Repeat; Lipoyl;
KW Complete proteome.
FT BINDING 41 41 LIPOYL (BY SIMILARITY).
FT ACT_SITE 159 159 LIPOYL (BY SIMILARITY).
FT CONFLICT 225 225 POTENTIAL.
FT CONFLICT 295 301 A -> V (IN REF. 1).
FT CONFLICT 328 329 GGAGAG -> AVAPR (IN REF. 1).
SQ SEQUENCE 547 AA; 56709 MW; 24E15CC9A590CB4 CRC64;
Query Match 33.8%; Score 115.5; DB 1; Length 547;
Best Local Similarity 37.5%; Pred. No. 0.00022;
Matches 27; Conservative 14; Mismatches 26; Indels 5; Gaps 1;
QY 3 EIPAPLAGTVSK-----ILVKGDTVAKGTVLVLEAMKMETEINAPTDGKVEKLVKER 57
Db 121 DIKVPDGSAGKANVIEVMKAGPTVEADSLTLESKASMEIPSPASGVESISIVG 180
QY 58 DAVGGGGLIKI 69
Db 181 DEVGTGLILKLT 192
ID BMB7_MYCTU STANDARD; PRT; 70 AA.
BMB7_MYCTU
AC 005845;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Biotinylated protein TB7.3.
GN RV3221C OR MT3317 OR MYCY07D11.05.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinobacteriales;
OC Actinomycetales; Corynebacteriaceae; Mycobacterium.
NCBI_TaxID=1773;
OX
RN
RP
RC STRAIN-H37RV;
RA MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Broesch R., Parkhill J., Garner T., Churcher C., Harris D.,
Gordon S.V., Eiglmeyer K., Gas S., Barry C.E. III, Tekala F.,
Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
Davies R., Devlin K., Fellwell T., Gentles S., Hamlin N., Holtroyd S.,
Hornsey T., Jagels K., Kirogh A., McLean J., Moule S., Murphy L.,
Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
Rutter S., Seeger K., Skelton S., Squares R.,
Stulston J.E., Taylor K., Whitehead S., Barrell B.G.,
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence."
RT Nature 393:537-544(1998).
RN
RP
RC STRAIN-CDC 1551 / Oshkosh;
RA Fietzmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
Kolony J.F., Nelson W.C., Umayan L.A., Ermolova M.D., Salzberg S.L.,
RA Delcher A., Utterback T., Weidman J., Khouli H., Gill J., Mikula A.,
RA Bishai W.;
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains.";

RL Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
RN
RP
RX MEDLINE=20072687; PubMed=10603390;
RA Skjot R.L., Oeltinger T., Rosenkrands I., Ravn P., Brock I.,
RA Jacobsen S., Andersen P.;
RT "Comparative evaluation of low-molecular-mass proteins from
RT Mycobacterium tuberculosis identifies members of the ESAT-6 family as
RT immunodominant T-cell antigens."
RL Infect. Immun. 68:214-220(2000).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: Z95120; CAB08316.1; -
DR EMBL: AE007143; AAK47659.1; ALT_INIT.
DR HSSP: P10802; 1LYU.
DR TIGR: MT3317; -
DR TUBERCULIST; RV3221c; -
DR InterPro: IPR001882; Biotin_attach.
DR InterPro: IPR000089; Biotin_lipoyl.
DR Pfam: PF00364; biotin_lipoyl; 1.
DR PROSITE: PS00189; BIOTIN; FALSE NEG.
KW Biotin; Antigen; Complete proteome.
FT INT_MET 0 0
FT BINDING 36 36 BIOTIN (BY SIMILARITY).
SQ SEQUENCE 70 AA; 7175 MW; 08B82DD3A76892D CRC64;
Query Match 33.6%; Score 115; DB 1; Length 70;
Best Local Similarity 37.7%; Pred. No. 3.1e-05;
Matches 23; Conservative 13; Mismatches 25; Indels 0; Gaps 0;
QY 3 EIPAPLAGTVSKIIIVKESDTVAKGTVLVLEAMKMETEINAPTDGKVEKLVKERDAVG 62
Db 3 DYRAEIVASVLEVYVNEEDQIDKGDVVLLSMKMEIVLVLEAGTVSKVAVSGDVIGA 62
QY 63 G 63
Db 63 G 63
ID ODP2_MYCCA STANDARD; PRT; 438 AA.
ODP2_MYCCA
AC 049110;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Dihydrolipoamide acetyltransferase component of pyruvate dehydrogenase
DE complex (EC 2.3.1.12) (E2).
GN PDHC OR ODP2.
OS Mycoplasma capricolum.
OC Bacteria; Firmicutes; Mollicutes; Entomoplasmatales;
OC Entomoplasmataceae.
NCBI_TaxID=2095;
OX
RN
RP
RC MEDLINE=97001869; PubMed=8844861;
RA Zhu P.P., Peterkofsky A.;
RT "Sequence and organization of genes encoding enzymes involved in
RT pyruvate metabolism in Mycoplasma capricolum."
RT Protein Sci. 5:1719-1736(1996).
CC
CC -1- FUNCTION: THE PYRUVATE DEHYDROGENASE COMPLEX CATALYZES THE OVERALL
CC CONVERSION OF PYRUVATE TO ACETYL-COA & CO(2). IT CONTAINS MULTIPLE
CC COPIES OF THREE ENZYMOLOGIC COMPONENTS: PYRUVATE DEHYDROGENASE (E1),
CC DIHYDROLIPOAMIDE ACETYLTRANSFERASE (E2) & LIPOAMIDE DEHYDROGENASE
CC (E3) (BY SIMILARITY).
CC
CC -1- CATALYTIC ACTIVITY: Acetyl-CoA + dihydrolipoamide -> CoA + S-

```
CC acetylhydroliponamide.
CC -1 COFACTOR: THE E2 COMPONENT CONTAINS ONE COVALENTLY-BOUND LIPOYL
CC CORRECTOR (BY SIMILARITY).
CC -1 SUBUNIT: FORMS A 24-POLYPEPTIDE STRUCTURAL CORE WITH OCTAHEDRAL
CC SYMMETRY (BY SIMILARITY).
CC -1 SIMILARITY: BELONGS TO THE 2-OXOACID DEHYDROGENASE FAMILY.
CC -1 SIMILARITY: CONTAINS 1 LIPOYL-BINDING DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U62057; AAC4344.1; -.
CC HSSP: P07016; 1C4T.
CC InterPro: IPR001078; 2oxoacid_dh.
CC InterPro: IPR000089; Biotin_lipoyl.
CC InterPro: IPR004167; E3_binding.
CC InterPro: IPR003016; Lipoyl.
CC Pfam: PF00198; 2-oxoacid_dh; 1.
CC Pfam: PF00364; biotin_lipoyl; 1.
CC Pfam: PF02817; e3_binding; 1.
CC ProDom: PD001115; 2oxoacid_dh; 1.
CC PROSITE: PS00189; LIPOYL; 1.
CC GLYCOSYL: Transferase: Acyltransferase: Lipoyl.
CC BINDING: 42 42 LIPOYL (BY SIMILARITY).
CC ACT_SITE: 411 411 POTENTIAL.
CC SEQUENCE: 438 AA; 46927 MW; 4BF83B697480B4AB CnC64;

Query Match 33.3%; Score 114; DB 1; Length 438;
Best Local Similarity 38.3%; Pred. NO. 0.00025;
Matches 23; Conservative 16; Mismatches 21; Indels 0; Gaps 0;

OY 10 GTVSTILVKEGPTVAKGTVLVLEAMKMETETINAPTDGKVEVLKENDAVOGGGLIKI 69
DB 16 GTVAELVAKGVGDKVKEGSLVFEIDKVNSELPAPVAGKIAVINIKAGEIIVGVVMEI 75

RESULT 28
ODP2_ECOLI STANDARD: PRT: 629 AA.
AC P06959;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Dihydroliponamide acetyltransferase component of pyruvate dehydrogenase
DE complex (EC 2.3.1.12) (E2).
GN ACET OR B0115.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-K12;
RC MEDLINE=83234434; PubMed=6345153;
RA Stephens P.E., Darlison M.G., Lewis H.M., Guest J.R.;
RT "The pyruvate dehydrogenase complex of Escherichia coli K12.
RT Nucleotide sequence encoding the dihydroliponamide acetyltransferase
RT component."
RT Eur. J. Biochem. 133:481-489(1983).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-K12 / MG1655;
RC MEDLINE=97426617; PubMed=9278503;
RA Blatter F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Coliaco-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
```

```
RL Science 277:1453-1474(1997).
RN [3]
RP SEQUENCE OF 1-12.
RC STRAIN-K12 / EMG2;
RC MEDLINE=97443975; PubMed=9298646;
RA Link A.J., Robison K., Church G.M.;
RT "Comparing the predicted and observed properties of proteins encoded
RT in the genome of Escherichia coli K-12."
RT Electrophoresis 18:1259-1313(1997).
RN [4]
RP SEQUENCE OF 34-46.
RC MEDLINE=84256520; PubMed=6821375;
RA Hale G., Perham R.N.;
RT "Amino acid sequence around lipotic acid residues in the pyruvate
RT dehydrogenase multienzyme complex of Escherichia coli."
RT Biochem. J. 187:905-908(1980).
RN [5]
RP MUTAGENESIS OF HIS-602.
RC MEDLINE=90351365; PubMed=2201286;
RA Russel G.C., Guest J.R.;
RT "Overexpression of restructured pyruvate dehydrogenase complexes and
RT site-directed mutagenesis of a potential active-site histidine
RT residue."
RT Biochem. J. 269:443-450(1990).
RN [6]
RP LIPOYLATED DOMAINS STUDIES.
RC MEDLINE=91024917; PubMed=2121129;
RA All S.T., Guest J.R.;
RT "Isolation and characterization of lipoylated and unlipooylated
RT domains of the E2 subunit of the pyruvate dehydrogenase complex of
RT Escherichia coli."
RT Biochem. J. 271:139-145(1990).
RN [7]
RP FUNCTION: The pyruvate dehydrogenase complex catalyzes the overall
RP conversion of pyruvate to acetyl-CoA and CO(2). It contains
RP multiple copies of three enzymatic components: pyruvate
RP dehydrogenase (E1), dihydroliponamide acetyltransferase (E2) and
RP liponamide dehydrogenase (E3).
CC -1 CATALYTIC ACTIVITY: Acetyl-CoA + dihydroliponamide -> CoA + S-
CC acetylhydroliponamide.
CC -1 COFACTOR: THE E2 COMPONENT CONTAINS THREE COVALENTLY-BOUND LIPOYL
CC COFACTORS.
CC -1 SUBUNIT: FORMS A 24-POLYPEPTIDE STRUCTURAL CORE WITH OCTAHEDRAL
CC SYMMETRY.
CC -1 SIMILARITY: BELONGS TO THE 2-OXOACID DEHYDROGENASE FAMILY.
CC -1 SIMILARITY: CONTAINS 3 LIPOYL-BINDING DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: V01498; CAA24741.1; -.
CC EMBL: D26562; BAA05673.1; -.
CC EMBL: AE000120; AAC73226.1; -.
CC PIR: A30278; XXCCDP.
CC PIR: A16026; A16026.
CC PIR: S45194; S45194.
CC HSSP: P10802; IDPC.
CC SWISS-2PAGE: P06959; COLI.
CC EC02DBASE: C062.7; 6TH EDITION.
CC EC02DBASE: C070.0; 6TH EDITION.
CC Ecogene: EG10025; acef.
CC InterPro: IPR001078; 2oxoacid_dh.
CC InterPro: IPR000089; Biotin_lipoyl.
CC InterPro: IPR004167; E3_binding.
CC InterPro: IPR003016; Lipoyl.
CC Pfam: PF00198; 2-oxoacid_dh; 1.
CC Pfam: PF00364; biotin_lipoyl; 3.
CC Pfam: PF02817; e3_binding; 1.
CC ProDom: PD001115; 2oxoacid_dh; 1.
```

DR * PROSITE; PS00189; LIPOYL; 3.
 KW Glycolysis; Transferrase; Acyltransferase; Repeat; Lipoyl;
 KM Complete proteome.
 FT INIT MET 0 0
 FT DOMAIN 1 314 LIPOYL BINDING, ACIDIC.
 FT DOMAIN 316 629 SUBUNIT BINDING, CATALYTIC.
 FT DOMAIN 372 388 HYDROPHOBIC.
 FT DOMAIN 541 566 HYDROPHOBIC.
 FT BINDING 40 40 LIPOYL.
 FT BINDING 143 143 LIPOYL.
 FT BINDING 244 244 LIPOYL.
 FT ACT_SITE 546 546 POTENTIAL.
 FT ACT_SITE 602 602 POTENTIAL.
 FT REPEAT 1 102 POTENTIAL.
 FT REPEAT 103 203
 FT REPEAT 204 313
 FT MUTAGEN 602 602 H->C: ABOLISHES CATALYTIC ACTIVITY.
 SQ SEQUENCE 629 AA; 65964 MW; 058751268B2CCCCC CRC64;
 Query Match 33.2%; Score 113.5; DB 1; Length 629;
 Best Local Similarity 34.0%; Pred. No. 0.0004;
 Matches 32; Conservative 9; Mismatches 26; Indels 27; Gaps 3;
 OY 3 EIPAPLAGTAVSKILVKEGDTVAKGQTVLVE-----AMKMET----- 39
 DB 147 EYAPAFAGTAVKKEIKVWDKSTGLIMFEVAGAGAAAPAKQAPAPAPAPAGVK 206
 OY 40 EINAPPTD-----KVEKVIYKERDANVGSGGLIKI 69
 DB 207 EVNVPDGGDEVEYEVKVGDKVAALAEQSLTV 240
 RESULT 29
 PCCA_RAT STANDARD: PRT: 704 AA.
 AC P14882;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-DEC-1992 (Rel. 24, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Propionyl-CoA carboxylase alpha chain, mitochondrial precursor
 DE (EC 6.4.1.3) (PCCase alpha subunit) (Propanoyl-CoA:carbon dioxide
 DE ligase alpha subunit) (Fragment).
 GN PCCA.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_Taxid=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-89308706; PubMed-2745462;
 RA Browner M.F., Taroni F., Szul E., Rosenberg L.E.;
 RT "Sequence analysis, biogenesis, and mitochondrial import of the
 RT alpha-subunit of rat liver propionyl-CoA carboxylase.";
 RL J. Biol. Chem. 264:12680-12685(1989).
 RN [2]
 RP REVISIONS.
 RA Browner M.F., Taroni F., Szul E., Rosenberg L.E.;
 RL Submitted (FEB-1989) to the EMBL/Genbank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: ATP + propanoyl-CoA + HCO(3) (-) = ADP +
 CC phosphate + (S)-methylmalonyl-CoA.
 CC -1- COFACTOR: BIOTIN.
 CC -1- PATHWAY: KEY ENZYME IN THE CATABOLIC PATHWAY OF ODD-CHAIN
 CC FATTY ACIDS, ISOLEUCINE, THREONINE, METHIONINE, AND VALINE.
 CC -1- SUBUNIT: PROBABLY A DODECAMER COMPOSED OF SIX BIOTIN-CONTAINING
 CC ALPHA SUBUNITS AND SIX BETA SUBUNITS.
 CC -1- SUBCELLULAR LOCATION: Mitochondrial matrix.
 CC -1- DISEASE: PROPIONIC ACIDEMIA DUE TO RECESSIVELY INHERITED
 CC DEFICIENCY OF PCCASE ACTIVITY OFTEN CAUSES LIFE-THREATENING
 CC KETOSIS AND ACIDOSIS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>
 CC or send an email to license@sdb-sdb.ch).
 CC -----
 CC DR EMBL; M22631; AAA88512.1; ALT_SEQ.
 CC DR PIR; A34337; A34337.
 CC DR HSP; P24182; IDV1.
 CC DR InterPro; IPR001882; Biotin_attach.
 CC DR InterPro; IPR000089; Biotin_lipoyl.
 CC DR InterPro; IPR000901; CPSase.
 CC DR Pfam; PF00289; CPSase_L_chain; 1.
 CC DR Pfam; PF00364; Biotin_lipoyl; 1.
 CC DR Pfam; PF02785; Biotin_carb_C; 1.
 CC DR Pfam; PF02786; CPSase_L_D2; 1.
 CC DR PROSITE; PS00188; BIOTIN; 1.
 CC DR PROSITE; PS00866; CPSASE_1; 1.
 CC DR PROSITE; PS00867; CPSASE_2; 1.
 CC DR Mitochondrion; Ligase; Biotin; ATP-binding; Transit peptide.
 CC FT NON_TER 1 1
 CC FT TRANSIT 1 21 MITOCHONDRION.
 CC FT CHAIN 22 704 PROPIONYL-CoA CARBOXYLASE ALPHA CHAIN.
 CC FT NP_BIND 199 204 ATP (POTENTIAL).
 CC FT ACT_SITE 329 329 BY SIMILARITY.
 CC FT BINDING 670 670 BIOTIN (BY SIMILARITY).
 CC SQ SEQUENCE 704 AA; 77711 MW; 36CEEC52DF2D2A8A CRC64;
 Query Match 32.5%; Score 111; DB 1; Length 704;
 Best Local Similarity 40.6%; Pred. No. 0.00081;
 Matches 26; Conservative 10; Mismatches 28; Indels 0; Gaps 0;
 OY 6 APLAGTAVSKILVKEGDTVAKGQTVLVEAMKMETEINAPTDKVEVYKERDANVGSGG 65
 DB 640 SPAPGVAVASVAPGDMVAGEDEICVIEAMKQNSMTAGKMGVKKLVHCKAGTVEGGL 699
 OY 66 LIKI 69
 DB 700 LEVEL 703
 RESULT 30
 PCCA_HUMAN STANDARD: PRT: 703 AA.
 ID P05165; 015979;
 AC P05165; 015979;
 DT 13-AUG-1987 (Rel. 05, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Propionyl-CoA carboxylase alpha chain, mitochondrial precursor
 DE (EC 6.4.1.3) (PCCase alpha subunit) (Propanoyl-CoA:carbon dioxide
 DE ligase alpha subunit).
 GN PCCA.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-89296507; PubMed-2740237;
 RA Lammonwah A.-M., Mahuran D.J., Gravel R.A.;
 RT "Human mitochondrial propionyl-CoA carboxylase: localization of the
 RT N-terminus of the pro- and mature alpha chains in the deduced
 RT primary sequence of a full-length cDNA.";
 RL Nucleic Acids Res. 17:4396-4396(1989).
 RN [2]
 RP REVISIONS.
 RA Gravel R.;
 RL Submitted (APR-1993) to the EMBL/Genbank/DBJ databases.
 CC [3]
 CC SEQUENCE FROM N.A.
 CC TISSUE-Placenta;
 CC Strausberg R.;
 RL Submitted (NOV-2000) to the EMBL/Genbank/DBJ databases.

Thu May 1 08:55:12 2003

us-09-987-485-2.rsp

Page 21

Search completed: May 1, 2003, 07:52:46
Job time : 7.19792 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4_p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using SW model

Run on: May 1, 2003, 07:49:48 ; Search time 38.2812 Seconds
(without alignments)
376.772 Million cell updates/sec

Title: US-09-987-485-2

Perfect score: 342
Sequence: 1 EGELPAPLACTSKVILKEG.....KVLKERDAVGGGGLIKIG 70

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 100 summaries

Database :

SPTREMBL_21: *
1: sp_archaea: *
2: sp_bacteria: *
3: sp_fungi: *
4: sp_human: *
5: sp_invertebrate: *
6: sp_mammal: *
7: sp_mmc: *
8: sp_organelle: *
9: sp_phage: *
10: sp_plant: *
11: sp_rodent: *
12: sp_virus: *
13: sp_vertebrate: *
14: sp_unclassified: *
15: sp_virus: *
16: sp_bacteriophage: *
17: sp_archaeophages: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	195	57.0	145	17	Q9V0A6
2	189	55.3	144	17	059021
3	181	52.9	144	17	08U303
4	180	52.6	144	17	08U303
5	178	52.0	140	17	028067
6	174	50.9	655	16	067484
7	174	50.9	984	2	09XBJ1
8	170	49.7	186	17	097YV7
9	167	48.8	134	16	09WZM6
10	167	48.8	596	2	048826
11	167	48.8	1144	16	097FR7
12	164	48.0	597	16	09PDM1
13	163	47.7	599	16	09PDM1
14	161	47.1	142	17	028194
15	159.5	46.6	576	16	08U917
16	159	46.5	665	16	092H13

17	158	46.2	169	17	Q97AR8	Q97AR8 sulfolobus
18	158	46.2	607	16	Q9HND1	Q9HND1 pseudomonas
19	157	45.9	1146	16	Q92CW1	Q92CW1 listeria in
20	157	45.9	1146	16	08Y846	08Y846 listeria mo
21	157	45.9	1150	16	09K9M0	09K9M0 bacillus ha
22	156	45.6	116	16	099ZL6	099ZL6 streptococ
23	155.5	45.5	1139	2	08ROL2	08ROL2 corynebacte
24	155.5	45.5	1140	2	054587	054587 corynebacte
25	153	44.7	135	16	08RAJ2	08RAJ2 thermomane
26	153	44.7	620	16	067544	067544 aquifex aeo
27	152	44.4	167	1	052603	052603 sulfolobus
28	152	44.4	436	10	040121	040121 lycopersico
29	151	44.2	166	16	099YD8	099YD8 streptococ
30	150	43.9	70	16	09KDS8	09KDS8 bacillus ha
31	150	43.9	122	16	08R7M0	08R7M0 thermomane
32	150	43.9	665	16	09ZCU3	09ZCU3 rickettsia
33	148	43.3	591	16	08XGK8	08XGK8 salmonella
34	148	43.3	1174	16	08UBX3	08UBX3 agrobacteri
35	147	43.0	132	16	099ZL1	099ZL1 streptococ
36	146	42.7	1150	16	099UY8	099UY8 staphylococ
37	145	42.4	655	16	091299	091299 pseudomonas
38	145	42.4	1147	2	P944A8	P944A8 bacillus st
39	144	42.1	161	16	09FBC1	09FBC1 streptococ
40	144	42.1	1137	2	09RAT6	09RAT6 lactococcus
41	144	42.1	1192	3	09HES8	09HES8 aspergillus
42	143.5	42.0	1124	16	09ARK6	09ARK6 streptomyc
43	143	41.8	145	2	09ZAA7	09ZAA7 acidimnoco
44	143	41.8	1137	16	09CHD7	09CHD7 lactococcus
45	143	41.8	1185	3	P78822	P78822 schizosacch
46	143	41.8	1185	3	09U0E1	09U0E1 schizosacch
47	142	41.5	155	16	09CHS5	09CHS5 lactococcus
48	142	41.5	602	16	09CL25	09CL25 pasteurella
49	141	41.2	1078	16	09A3J0	09A3J0 caulobacter
50	140	40.9	1152	16	098F27	098F27 rhizobium l
51	140	40.9	1175	3	08XIT3	08XIT3 pichia angu
52	140	40.9	1193	3	093918	093918 aspergillus
53	139.5	40.8	1127	16	P95127	P95127 mycobacteri
54	139	40.6	134	16	08R5Y8	08R5Y8 fuscobacteri
55	138.5	40.5	1127	2	09F843	09F843 mycobacteri
56	138	40.4	129	2	057111	057111 yellowella
57	137	40.1	601	2	09RFP9	09RFP9 haemophilus
58	136	39.8	134	2	054030	054030 propionigen
59	136	39.8	598	2	08RNO3	08RNO3 amycolatops
60	136	39.8	1152	16	092L13	092L13 rhizobium m
61	135	39.8	1154	2	059740	059740 rhizobium e
62	135	39.5	661	16	09HZV6	09HZV6 pseudomonas
63	134	39.2	1180	13	09DPT1	09DPT1 brachydanio
64	133.5	39.0	159	16	097DA8	097DA8 clostridium
65	132.5	38.7	162	16	08UFR6	08UFR6 agrobacteri
66	132.5	38.7	170	16	09A743	09A743 caulobacter
67	131.5	38.5	158	16	0920K0	0920K0 rhizobium m
68	130.5	38.2	654	16	09A6C6	09A6C6 caulobacter
69	129	37.7	203	10	08WZC0	08WZC0 oryza sativ
70	129	37.7	591	2	08R0M4	08R0M4 corynebacte
71	129	37.7	667	16	P71538	P71538 mycobacteri
72	129	37.7	920	2	09KW58	09KW58 bacillus su
73	129	37.7	935	11	062043	062043 mus musculu
74	129	37.7	1148	16	09KWT4	09KWT4 bacillus su
75	128.5	37.6	155	16	08XVP4	08XVP4 ralsionia s
76	128.5	37.6	593	16	083095	083095 treponema p
77	128.5	37.6	610	17	09HPB8	09HPB8 halobacteri
78	127	37.1	600	16	P96890	P96890 mycobacteri
79	126	36.8	591	2	P71122	P71122 corynebacte
80	126	36.8	678	16	08YF3	08YF3 bruceella me
81	125.5	36.7	152	16	08XRB7	08XRB7 ralsionia s
82	125	36.5	712	16	08Y2S0	08Y2S0 ralsionia s
83	124.5	36.4	156	16	08Y2S0	08Y2S0 ralsionia s
84	124.5	36.4	1181	5	09XZ00	09XZ00 drosophila
85	124.5	36.4	1196	5	0917E9	0917E9 drosophila
86	124	36.3	665	5	060966	060966 leishmania
87	123	36.0	456	2	09LBC6	09LBC6 rhodospirill
88	122	35.7	572	17	08TSX1	08TSX1 methanosarc
89	122	35.7	694	16	0985D4	0985D4 rhizobium l

90 121.5 35.5 1124 2 050450
 91 120.7 35.4 1207 16 094797
 92 120.5 35.2 187 16 094797
 93 120 35.1 724 5 019842
 94 119.5 34.9 154 16 094797
 95 119.5 34.9 158 2 054761
 96 119 34.8 573 1 094797
 97 119 34.8 584 2 094797
 98 119 34.8 591 2 054105
 99 119 34.8 597 2 054105
 100 119 34.8 1158 16 087320

ALIGNMENTS

RESULT 1
 09Y0A6 PRELIMINARY: PRT: 145 AA.
 ID 09Y0A6
 AC 09Y0A6
 DT 01-MAY-2000 (TREMblrel. 13, Created)
 DT 01-MAY-2000 (TREMblrel. 13, last sequence update)
 DT 01-MAR-2002 (TREMblrel. 20, last annotation update)
 DE Methylmalonyl-CoA decarboxylase gamma chain.
 GN PAB1771.
 OS Pyrococcus abyssi.
 OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
 OC Pyrococcus.
 OX NCBI_TaxID=29292;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN-ORSAV;
 RA Heilig R.;
 RT "Pyrococcus abyssi genome sequence: Insights into archaeal chromosome structure and evolution."
 RL Submitted (Jul-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: A248285; CAB49799.1;
 DR HSSP: P02905; IBD0.
 DR InterPro: IPR001882; Biotin_attach.
 DR InterPro: IPR000089; Biotin_lipoyl.
 DR Pfam: PF00364; biotin_lipoyl.1.
 DR PROSITE: PS00188; BIOTIN; UNKNOWN_1.
 KW Complete proteome.
 SQ SEQUENCE 145 AA: 15489 MW: 9C1443363F40D94 CRC64;
 Query Match 57.0%; Score 195; DB 17; Length 145;
 Best Local Similarity 58.6%; Pred. No. 1.6e-12;
 Matches 41; Conservative 10; Mismatches 19; Indels 0; Gaps 0;
 QY 1 EGTTPPLAGTYSKIIYKESDVTYKAGQVLYLEAMKMEETINAPTDGKYEKLVKERDAV 60
 DB 76 ENVTAPMPGKVLKILVQEGQVTLGCGILLIEAMKMEETINAPTDGKYEKLVKERDAV 135
 QY 61 QGGGGLIKIG 70
 DB 136 DTGTPLIEIG 145
 RESULT 2
 059021 PRELIMINARY: PRT: 149 AA.
 ID 059021
 AC 059021
 DT 01-AUG-1998 (TREMblrel. 07, Created)
 DT 01-AUG-1998 (TREMblrel. 07, last sequence update)
 DT 01-MAR-2002 (TREMblrel. 20, last annotation update)
 DE 149AA long hypothetical methylmalonyl-CoA decarboxylase gamma chain.
 GN PH1284.
 OS Pyrococcus horikoshii.
 OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
 OC Pyrococcus.
 OX NCBI_TaxID=53953;
 RN (1)
 RP SEQUENCE FROM N.A.

RC STRAIN-OR3;
 RX MEDLINE=98344137; PubMed=9679194;
 RA Kawarabayashi Y., Sawada M., Horikawa H., Halkawa Y., Hino Y.,
 RA Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai Y.,
 RA Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohfuku Y.,
 RA Funahashi T., Tanaka T., Kudo H., Yamazaki J., Kishida N., Oguchi A.,
 RA Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,
 RA Masuchi Y., Shizuya H., Kikuchi H.;
 RT "Complete sequence and gene organization of the genome of a hyper-
 thermophilic archaeobacterium, Pyrococcus horikoshii 073.";
 RL DNA Res. 5:55-76(1998).
 DR EMBL: AP000005; BAA30387.1;
 DR HSSP: P10802; I110.
 DR InterPro: IPR001882; Biotin_attach.
 DR InterPro: IPR000089; Biotin_lipoyl.
 DR Pfam: PF00364; biotin_lipoyl.1.
 DR PROSITE: PS00188; BIOTIN; UNKNOWN_1.
 KW Complete proteome.
 SQ SEQUENCE 149 AA: 15985 MW: 1C3A5F47E6BA6F1 CRC64;

Query Match 55.3%; Score 189; DB 17; Length 149;
 Best Local Similarity 54.3%; Pred. No. 6.7e-12;
 Matches 38; Conservative 14; Mismatches 18; Indels 0; Gaps 0;

QY 1 EGTTPPLAGTYSKIIYKESDVTYKAGQVLYLEAMKMEETINAPTDGKYEKLVKERDAV 60
 DB 80 ENVTAPMPGKVLKILVQEGQVTLGCGILLIEAMKMEETINAPTDGKYEKLVKERDAV 139
 QY 61 QGGGGLIKIG 70
 DB 140 DTGTPLIEIG 149

RESULT 3
 08U303 PRELIMINARY: PRT: 144 AA.
 ID 08U303
 AC 08U303
 DT 01-JUN-2002 (TREMblrel. 21, Created)
 DT 01-JUN-2002 (TREMblrel. 21, last sequence update)
 DT 01-JUN-2002 (TREMblrel. 21, last annotation update)
 DE Methylmalonyl-CoA decarboxylase gamma chain.
 GN PF0673.
 OS Pyrococcus furiosus.
 OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
 OC Pyrococcus.
 OX NCBI_TaxID=2261;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN-VCI / DSM 3638 / ATCC 43587 / JCM 8422;
 RA Weiss R.B., Dunn D.M., Robb F.T., Brown J.R.;
 RT "The complete sequence of the Pyrococcus furiosus genome."
 RL Submitted (Feb-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AE010188; AAL80797.1;
 KW Complete proteome.
 SQ SEQUENCE 144 AA: 15315 MW: 422C96A8ED809C6A CRC64;

Query Match 52.9%; Score 181; DB 17; Length 144;
 Best Local Similarity 55.2%; Pred. No. 4.2e-11;
 Matches 37; Conservative 10; Mismatches 20; Indels 0; Gaps 0;

QY 4 ITPPLAGTYSKIIYKESDVTYKAGQVLYLEAMKMEETINAPTDGKYEKLVKERDAV 63
 DB 78 VTPAPMPGKVLKILVQEGQVTLGCGILLIEAMKMEETINAPTDGKYEKLVKERDAV 137
 QY 64 QGLIKIG 70
 DB 138 QPLIEIG 144

RESULT 4
 058564 PRELIMINARY: PRT: 571 AA.
 ID 058564
 AC 058564

DT 01-AUG-1998 (TREMBLrel. 07, Created)
 DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE 571AA long hypothetical oxaloacetate decarboxylase alpha chain.
 GN PH0834.
 OS Pyrococcus horikoshii.
 OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
 OC Pyrococcus.
 OX NCBI_TaxID=53953;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=OT3;
 RX MEDLINE=96344137; PubMed=9679194;
 RA Kawarabayashi Y., Sawada M., Horikawa H., Halkawa Y., Hino Y.,
 RA Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai Y.,
 RA Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohfuku Y.,
 RA Funahashi T., Tanaka T., Kudoh Y., Yamazaki J., Kushida N., Oguchi A.,
 RA Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,
 RA Masuchi Y., Shizuya H., Kikuchi H.;
 RT "Complete sequence and gene organization of the genome of a hyper-
 thermophilic archaeobacterium, Pyrococcus horikoshii OT3.";
 RL DNA Res. 5:55-76(1998).
 DR EMBL: AP000003; BAA2928.1;
 DR HSSP: P02905; IBD0.
 DR InterPro: IPR001882; Biotin_attach.
 DR InterPro: IPR000089; Biotin_lipoyl.
 DR InterPro: IPR000891; HMGL-like.
 DR InterPro: IPR003379; PTC_ODA.
 DR Pfam: PF00364; biotin_lipoyl; 1.
 DR Pfam: PF00682; HMGL-like; 1.
 DR Pfam: PF02436; PTC_ODA; 1.
 DR TIGRFAMs: TIGR01108; cda; 1.
 DR PROSITE: PS00188; BIOTIN; UNKNOWN_1.
 KW Complete proteome.
 SQ SEQUENCE 571 AA; 64748 MW; B308B0B1C0E5B103 CRC64;
 Query Match 52.6%; Score 180; DB 17; Length 571;
 Best Local Similarity 54.5%; Pred. No. 2.5e-10;
 Matches 36; Conservative 14; Mismatches 16; Indels 0; Gaps 0;
 QY 4 IPAPLAGVSKILVKEGDIYKAGQVTVLLEAMKMETETNAPDGVKVEKLVKERNDAVCG 63
 DB 505 VSAPMPGVKLVKRVGRVRRVGGGLVLEAMKMENTIPSPDGVKRLVKEGAVDVG 564
 QY 64 QGLIKI 69
 DB 565 QPLIEL 570
 RESULT 5
 O28067 PRELIMINARY; PRT; 140 AA.
 AC O28067;
 DT 01-JAN-1998 (TREMBLrel. 05, Created)
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Methylenalonyl-CoA decarboxylase, biotin carboxyl carrier subunit
 DE (MADC).
 GN AF2216.
 OS Archaeoglobus fulgidus.
 OC Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;
 OC Archaeoglobaceae; Archaeoglobus.
 OX NCBI_TaxID=2234;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
 RX MEDLINE=96049343; PubMed=9389475;
 RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
 RA Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.R., Peterson J.D.,
 RA Richardson D.L., Kerlavage A.R., Graham D.E., Kyriakides N.C.,
 RA Fleischman R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
 RA Kitzness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
 RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,

RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,
 RA Cotton M.D., Spriggs T., Artlich P., Kaine B.P., Sykes S.M.,
 RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
 RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
 RA Venter J.C.;
 RT "The complete genome sequence of the hyperthermophilic, sulphate-
 reducing archaeon Archaeoglobus fulgidus.";
 RL Nature 390:364-370(1997).
 DR EMBL: AE000952; AAB89036.1;
 DR HSSP: P20708; IGH1.
 DR TIGR: AF2216;
 DR InterPro: IPR001882; Biotin_attach.
 DR InterPro: IPR000089; Biotin_lipoyl.
 DR Pfam: PF00364; biotin_lipoyl; 1.
 DR PROSITE: PS00188; BIOTIN; UNKNOWN_1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 140 AA; 15686 MW; 30B449C45489C14A CRC64;
 Query Match 52.0%; Score 178; DB 17; Length 140;
 Best Local Similarity 54.7%; Pred. No. 8.4e-11;
 Matches 35; Conservative 13; Mismatches 16; Indels 0; Gaps 0;
 QY 4 IPAPLAGVSKILVKEGDIYKAGQVTVLLEAMKMETETNAPDGVKVEKLVKERNDAVCG 63
 DB 74 IPAPAGVTKILKRVGEKAGETVLLLEAMKMENTIPSPDGVKRLVKEGDIYKAGV 133
 QY 64 QGLI 67
 DB 134 DVLV 137
 RESULT 6
 O67484 PRELIMINARY; PRT; 655 AA.
 AC O67484;
 DT 01-AUG-1998 (TREMBLrel. 07, Created)
 DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Pyruvate carboxylase C-terminal domain.
 GN PYCA OR AO.1520.
 OS Aquifex aeolicus.
 OC Bacteria; Aquificae; Aquificales; Aquificaceae; Aquifex.
 OX NCBI_TaxID=63363;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=VF5.
 RX MEDLINE=98196666; PubMed=9537320;
 RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
 RA Graham D.E., Overbeek R., Shead M.A., Keller M., AuJay M., Huber R.,
 RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
 RT "The complete genome of the hyperthermophilic bacterium Aquifex
 aeolicus.";
 RL Nature 392:353-358(1998).
 DR EMBL: AE000744; AAC07445.1;
 DR HSSP: P02905; IBD0.
 DR InterPro: IPR001882; Biotin_attach.
 DR InterPro: IPR000089; Biotin_lipoyl.
 DR InterPro: IPR003379; PTC_ODA.
 DR Pfam: PF00364; biotin_lipoyl; 1.
 DR Pfam: PF00682; HMGL-like; 1.
 DR Pfam: PF02436; PTC_ODA; 1.
 DR PROSITE: PS00188; BIOTIN; UNKNOWN_1.
 KW Pyruvate; Complete proteome.
 SQ SEQUENCE 655 AA; 73612 MW; 2839436F6FED05D6 CRC64;
 Query Match 50.9%; Score 174; DB 16; Length 655;
 Best Local Similarity 50.0%; Pred. No. 1.2e-09;
 Matches 34; Conservative 13; Mismatches 21; Indels 0; Gaps 0;
 QY 2 GEIPAPLAGVSKILVKEGDIYKAGQVTVLLEAMKMETETNAPDGVKVEKLVKERNDAV 61
 DB 578 GVTPMPGVKLVKRVGRVRRVGGGLVLEAMKMENTIPSPDGVKRLVKEGAVDVG 637

```

QY 62 GGGGLIKI 69
Db 638 PDQAIMRI 645

RESULT 7
09XBJ1 PRELIMINARY; PRT; 984 AA.
AC 09XBJ1;
ID 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Pyruvate carboxylase (EC 6.4.1.1) (Fragment).
GN PCA.
OS Bacillus cereus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Bacillaceae; Bacillus.
OX NCBI_TaxID=1396;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 10987;
RX MEDLINE=99231848; PubMed=10217496;
RA Ostad O.A., Hegna I., Lindbeck T., Rishovd A.L., Kolsto A.B.;
RT "Genome organisation is not conserved between Bacillus cereus and
RT Bacillus subtilis.";
RL Microbiology 145:621-631(1999).
EMBL: AJ010111; CAB40604.1; -.
DR HSSP; P24182; IDV2.
DR Interpro: IPR000089; Biotin_lipoyl.
DR Interpro: IPR000091; CPase.
DR Interpro: IPR003379; PYC-OADA.
DR Pfam: PF02785; Biotin_carb_C; 1.
DR Pfam: PF00364; Biotin_lipoyl; 1.
DR Pfam: PF02786; CPase_L_D2; 1.
DR Pfam: PF00682; HMG_Like; 1.
DR Pfam: PF02436; PYC-OADA; 1.
DR PRINTS: PRO0098; CPASE.
DR TIGRfams: TIGR01235; pyruv_carbox; 1.
DR PROSITE: PS00867; CPASE_2; UNKNOWN_1.
KW Ligase.
FT NON_TER
SQ SEQUENCE 984 AA; 110135 MW; 7AB52F8D453A147D CRC64;

Query Match 50.9%; Score 174; DB 2; Length 984;
Best Local Similarity 53.0%; Pred. No. 1.9e-09;
Matches 35; Conservative 13; Mismatches 18; Indels 0; Gaps 0;

QY 4 IPAPLAGVSKILVKEGDYKAGQTVLVLEAMKETEINAPTDGKYEKLVKERDVG 63
Db 917 ISATMGTYAKVYVKGDEYKGDSDMAITFAMKETTVAQFNGKVKYVNDGDAIQTG 976

QY 64 GGLIKI 69
Db 977 DLIEL 982

RESULT 8
097VY7 PRELIMINARY; PRT; 186 AA.
AC 097VY7;
ID 01-OCT-2001 (TREMBLrel. 18, Created)
DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)
DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Biotin carboxyl carrier protein of propionyl-CoA carboxylase beta
DE subunit (EC 6.4.1.3).
GN SSO2464.
OS Sulfolobus solfataricus.
OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
OC Sulfolobus.
OX NCBI_TaxID=2287;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 35092 / DSM 1617 / P2;
RX MEDLINE=21332296; PubMed=11427726;
RA She O., Singh R.K., Confalonieri F., Zivanovic Y., Allard G.,
RA Aways M.J., Chan-Weiher C.C.-Y., Clausen I.G., Curtis B.A.,
RA De Moers A., Erasus G., Fletcher C., Gordon P.M.K.,
RA Helkamp-de Jong I., Jeffries A.C., Kozera C.J., Medina N., Peng X.,
RA Thi-Ngoc H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N.,
RA Charlebois R.L., Doolittle W.F., Duguet M., Gaasterland T.,
RA Garrett R.A., Ragan M.A., Sensen C.W., Van der Oost J.;
RT "The complete genome of the crenarchaeon Sulfolobus solfataricus P2.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).
DR EMBL: AE006845; AAK42603.1; -.
DR Interpro: IPR001882; Biotin_attach.
DR Interpro: IPR000089; Biotin_lipoyl.
DR Pfam: PF00364; Biotin_lipoyl; 1.
DR PROSITE: PS00188; Biotin; UNKNOWN_1.
KW ligase; Complete proteome.
SQ SEQUENCE 186 AA; 21166 MW; DB26587C39883B08 CRC64;

Query Match 49.7%; Score 170; DB 17; Length 186;
Best Local Similarity 55.1%; Pred. No. 7.6e-10;
Matches 38; Conservative 9; Mismatches 22; Indels 0; Gaps 0;

QY 1 EGEPAPLACTVSKILVKEGDYKAGQTVLVLEAMKETEINAPTDGKYEKLVKERDVG 60
Db 117 EGELVSPLEFGVYKIRYKEDAVNKGQPLISIDAMKETVISPGLGVQKILKEGGV 176

QY 61 GGGGLIKI 69
Db 177 KKGDILVI 185

RESULT 9
09WZH6 PRELIMINARY; PRT; 134 AA.
AC 09WZH6;
ID 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DE 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Propionyl-CoA carboxylase, gamma subunit.
GN TM0717.
OS Thermotoga maritima.
OC Bacteria; Thermotogae; Thermotogales; Thermotogaceae; Thermotoga.
OX NCBI_TaxID=2336;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MSB / DSM 3109;
RX MEDLINE=99287316; PubMed=10360571;
RA Nelson K.E., Clayton R.A., Gill S.R., Gwinn M.L., Dodson R.J.,
RA Haft D.H., Hickey E.R., Peterson J.D., Nelson W.C., Ketchum K.A.,
RA McDonald L., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M.,
RA Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
RA Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
RA Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
RT "Evidence for lateral gene transfer between Archaea and Bacteria from
RT genome sequence of Thermotoga maritima.";
RL Nature 399:323-329(1999).
DR EMBL: AE001743; AAD35799.1; -.
DR HSSP: P02905; LBDO.
DR TIGR: TM0717; -.
DR Interpro: IPR000089; Biotin_lipoyl.
DR Pfam: PF00364; Biotin_lipoyl; 1.
KW Complete proteome.
SQ SEQUENCE 134 AA; 15131 MW; A3BFBAE8574EC1 CRC64;

Query Match 48.8%; Score 167; DB 16; Length 134;
Best Local Similarity 51.5%; Pred. No. 1.1e-09;
Matches 34; Conservative 12; Mismatches 20; Indels 0; Gaps 0;

QY 4 IPAPLAGVSKILVKEGDYKAGQTVLVLEAMKETEINAPTDGKYEKLVKERDVG 63
Db 68 VKAPMAGIVLAKVYVKGQVNVGDKLVLEAMKEMENLOSFGYKELVLEAGNIEFG 127

```

```

QY 64 GGLIKI 69
DB 128 QILMKI 133

RESULT 10
Q48826 PRELIMINARY; PRT; 596 AA.
AC Q48826;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Oxalacetate decarboxylase alpha-chain.
OS OADA.
OC Bacteria; Proteobacteria; gamma subdivision; Legionellaceae group;
OC Legionellaceae; Legionella.
ON NCBI_TaxID=446;
RX STRAIN-CORBY;
RP SEQUENCE FROM N.A.
RX MEDLINE=97120897; Pubmed=8961567;
RA Jain B., Brand B.C., Lueck P.C., Di Beraardino M., Dimroth P.,
RA Hacker J.;
RT "An oxalacetate decarboxylase homologue protein influences the
RT intracellular survival of Legionella pneumophila.";
RL FEMS Microbiol. Lett. 145:273-279(1996).
CC -1- COFACTOR: BIOTIN (BY SIMILARITY).
DR EMBL: X99678; CAA67994.1; -.
DR HSP: P11961; 11AB.
DR InterPro: IPR001882; Biotin_attach.
DR InterPro: IPR000891; HMG_L-like.
DR InterPro: IPR003379; HMG_L-like.
DR InterPro: IPR000634; S/T_dehydrtse.
DR Pfam: PF00364; Biotin_11poyl; 1.
DR Pfam: PF02436; HMG_L-like; 1.
DR TIGRfams: TIGR01108; Oada; 1.
DR PROSITE: PS00188; BIOTIN; 1.
DR PROSITE: PS00165; DEHYDRATASE_SER_THR; 1.
KW Biotin.
SQ SEQUENCE 596 AA; 65667 MW; 55DBEAF96919C86 CRC64;

Query Match 48.8%; Score 167; DB 2; Length 596;
Best Local Similarity 51.5%; Pred. No. 5.8e-09;
Matches 35; Conservative 12; Mismatches 21; Indels 0; Gaps 0;

QY 2 GEIPAPLAGTYSKILVKGDTYKAGQTVLVLEAMKMETEINAPTDGKYEKYLKERDAVQ 61
DB 526 GDIVIAIPGSIHIVHSGDEVKAGQAVLVLEAMKMETEIKAPANGVAHILCKGDKVT 585
QY 62 GGGGLIKI 69
DB 586 PGQVLIRV 593

RESULT 11
Q97FR7 PRELIMINARY; PRT; 1144 AA.
AC Q97FR7;
DT 01-OCT-2001 (TREMBLrel. 18, Created)
DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Pyruvate carboxylase, PYCA.
OS CAC2660.
OC Clostridium acetobutylicum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC Clostridiales; Clostridiaceae; Clostridium.
ON NCBI_TaxID=1488;
RX [1]
RP SEQUENCE FROM N.A.

```

```

RC STRAIN-ATCC 824 / DSM 792 / VKM B-1787;
RX MEDLINE=21359325; Pubmed=11466286;
RA Noelling J., Breton G., Omelchenko M.V., Makarova K.S., Zeng Q.,
RA Gibson R., Lee H.M., Dubois J., Qiu D., Hitti Y.I.,
RA Tatusov R.L., Sabathe F., Doucette-Stamm L., Souaille P., Daly M.J.,
RA Bennett G.N., Koonin E.V., Smith D.R.;
RT "Genome sequence and comparative analysis of the solvent-producing
RT bacterium Clostridium acetobutylicum.";
RL J. Bacteriol. 183:4823-4838(2001).
DR EMBL: AE007763; AAK80607.1; -.
DR InterPro: IPR001882; Biotin_attach.
DR InterPro: IPR000891; CPSase.
DR InterPro: IPR000891; HMG_L-like.
DR InterPro: IPR003379; PYC_OADA.
DR Pfam: PF02785; Biotin_card_C; 1.
DR Pfam: PF00364; Biotin_11poyl; 1.
DR Pfam: PF02889; CPSase_L_D2; 1.
DR Pfam: PF02786; CPSase_L_D2; 1.
DR Pfam: PF00682; HMG_L-like; 1.
DR Pfam: PF02436; PYC_OADA; 1.
DR PRINTS: PR00098; CPSASE.
DR SMART: SM00481; POLITING; 1.
DR TIGRfams: TIGR01235; PYRUV_carbox; 1.
DR PROSITE: PS00188; BIOTIN; UNKNOWN_1.
DR PROSITE: PS00866; CPSASE_1; 1.
DR PROSITE: PS00867; CPSASE_2; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 1144 AA; 127709 MW; 519FA29A8008F326 CRC64;

Query Match 48.8%; Score 167; DB 16; Length 1144;
Best Local Similarity 50.7%; Pred. No. 1.2e-08;
Matches 34; Conservative 11; Mismatches 22; Indels 0; Gaps 0;

QY 3 EIPAPLAGTYSKILVKGDTYKAGQTVLVLEAMKMETEINAPTDGKYEKYLKERDAVQ 62
DB 1077 EIGASIPGNVYKVPKDPKDKSDLVLEAMKMETVNSVSDGTGIVKEDQVOS 1136
QY 63 GGGGLIKI 69
DB 1137 GQLVKL 1143

RESULT 12
Q9KUH1 PRELIMINARY; PRT; 597 AA.
AC Q9KUH1;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Oxalacetate decarboxylase, alpha subunit.
CN VC0550.
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
ON NCBI_TaxID=666;
RX [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EL TOR N16961 / SEROTYPE O1;
RX MEDLINE=20406833; Pubmed=10952301;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragol I., Sellers P.,
RA McDonald L., Uitterback T., Fleischmann R.D., Nierman W.C., White O.,
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RA Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
RT cholerae.";
RL Nature 406:477-483(2000).
CC -1- COFACTOR: BIOTIN (BY SIMILARITY).
DR EMBL: AE004141; AAF93718.1; -.
DR HSP: P02905; IBDO.

```

```
DR TIGR: VC0550: -
DR InterPro: IPR001882; Biotin_attach.
DR InterPro: IPR000089; Biotin_lipoyl.
DR InterPro: IPR000891; HMGL-like.
DR InterPro: IPR003379; PYC-OADA.
DR Pfam: PF00364; biotin_lipoyl. 1.
DR Pfam: PF00682; HMGL-like. 1.
DR Pfam: PF02436; PYC-OADA. 1.
DR TIGRPFAMS: TIGR01108; oada. 1.
DR PROSITE: PS00188; BIOTIN. 1.
DR Biotin: Complete proteome.
SQ SEQUENCE 597 AA; 64795 MW; 033BF5F2209F5468 CRC64;

Query Match 48.0%; Score 164; DB 16; Length 597;
Best Local Similarity 50.0%; Pred. No. 1,2e-08;
Matches 33; Conservative 14; Mismatches 19; Indels 0; Gaps 0;

OY 4 IIPAPLAGTIVKEDGYKAGOTVLYLEAMKMETETINAPTDGKYEKVLKERDAY 63
DB 531 VAAPLACTIRKIVGQDEVAEGDVLIVLEAMKMETETIRARSGVIOELHVKEGDSVRVG 590

OY 64 QGLIKI 69
DB 591 ASLSTL 596

RESULT 13
O9PP00 PRELIMINARY; PRT; 599 AA.
ID O9PP00:
AC O9PP00:
DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
DE 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE Putative pyruvate carboxylase B subunit (EC 6.4.1.1).
GN PYCB OR CJO933C.
OS Campylobacter jejuni.
OC Bacteria; Proteobacteria; epsilon subdivision; Campylobacter group;
OC Campylobacter.
OX NCBI_TaxID=197;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN-NCTC 11168;
RX MEDLINE=20150912; PubMed=10688204;
RA Parkhill J., Wren B.W., Mungall K., Ketley J.M., Churcher C.,
RA Basham D., Chillingworth T., Davies R.M., Feltham T., Holtroyd S.,
RA Jagels K., Karlyshev A.V., Moule S., Pallen M.J., Penn C.W.,
RA Quail M.A., Rattandream M.A., Rutherford K.M., van Vliet A.H.M.,
RA Whitehead S., Barrell B.G.;
RT "The genome sequence of the food-borne pathogen Campylobacter jejuni
RT reveals hypervariable sequences.";
RL Nature 403:665-668(2000).
CC -1- COFACTOR: BIOTIN (BY SIMILARITY).
DR EMBL: AL139076; CAB73190.1; -.
DR HSSP: P20708; 1GHT.
DR InterPro: IPR001882; Biotin_attach.
DR InterPro: IPR000089; Biotin_lipoyl.
DR InterPro: IPR000891; HMGL-like.
DR Pfam: PF00364; biotin_lipoyl. 1.
DR Pfam: PF00682; HMGL-like. 1.
DR PROSITE: PS00188; BIOTIN. 1.
DR Biotin: Complete proteome.
SQ SEQUENCE 599 AA; 65833 MW; E5C075F11AE40444 CRC64;

Query Match 47.7%; Score 163; DB 16; Length 599;
Best Local Similarity 50.0%; Pred. No. 1,5e-08;
Matches 33; Conservative 13; Mismatches 20; Indels 0; Gaps 0;

OY 1 EGIIRPPLAGTIVKESDGYKAGOTVLYLEAMKMETETINAPTDGKYEKVLKERDAY 60
DB 529 ENEVLAGISGNVFKIYVNEBEEKVSGOAIMVLEAMKMETETIRARSGVIOELHVKEGDSVRVG 588

OY 61 QGGGGL 66
DB 591 ASLSTL 596

DB 589 NEGEVL 594

RESULT 14
O28194 PRELIMINARY; PRT; 142 AA.
ID O28194:
AC O28194:
DT 01-JAN-1998 (TREMblrel. 05, Created)
DT 01-JAN-1998 (TREMblrel. 05, Last sequence update)
DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)
DE Oxaloacetate decarboxylase, biotin carboxyl carrier subunit, putative.
GN AF2085.
OS Archaeoglobus fulgidus.
OC Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;
OC Archaeoglobaceae; Archaeoglobus.
OX NCBI_TaxID=2234;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
RX MEDLINE=98049343; PubMed=9389475;
RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
RA Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
RA Richardson D.L., Kervagge A.R., Graham D.E., Kyrpides N.C.,
RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
RA Peterson R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,
RA Cotton M.D., Spriggs T., Artlich P., Kaine B.P., Sykes S.M.,
RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA Venter J.C.;
RT "The complete genome sequence of the hyperthermophilic, sulphate-
RT reducing archaeon Archaeoglobus fulgidus.";
RL Nature 390:364-370(1997).
DR EMBL: AE000960; AAB89171.1; -.
DR HSSP: P10802; 1IYT.
DR TIGR: AF2085; -.
DR InterPro: IPR000089; Biotin_lipoyl.
DR Pfam: PF00364; biotin_lipoyl. 1.
DR Hypothetical protein; Complete proteome.
SQ SEQUENCE 142 AA; 15573 MW; 73439FBD49E469A1 CRC64;

Query Match 47.1%; Score 161; DB 17; Length 142;
Best Local Similarity 50.8%; Pred. No. 4,7e-09;
Matches 32; Conservative 13; Mismatches 18; Indels 0; Gaps 0;

OY 1 EGIIRPPLAGTIVKESDGYKAGOTVLYLEAMKMETETINAPTDGKYEKVLKERDAY 60
DB 70 ENAVTSMLEGVVLIKIVGDKYKAGEPVYVESKMEKETIVSPTEGVATILVKEGQRI 129

OY 61 QGG 63
DB 130 EAG 132

RESULT 15
O80917 PRELIMINARY; PRT; 576 AA.
ID O80917:
AC O80917:
DT 01-JUN-2002 (TREMblrel. 21, Created)
DT 01-JUN-2002 (TREMblrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE Biotin carboxylase.
GN ATU913 OR AGR L_1864.
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Rhizobium.
OX NCBI_TaxID=176299;
RN (1)
RP SEQUENCE FROM N.A.
RC MEDLINE=21608550; PubMed=11743193;
RX Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
```

RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Boyce D. Sr.,
 RA Chapman P., Clendinning J., Deacherae G., Gillet W., Grant C.,
 RA Kutayavin T., Levy R., Li M.-J., McLelland E., Palmieri A.,
 RA Raymond C., Rouse G., Saenphumachak C., Wu Z., Romero P., Gordon D.,
 RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
 RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
 RA Chumley P., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
 RA Nester E.W.,
 RT "The genome of the natural genetic engineer *Agrobacterium tumefaciens*
 RT C58."
 RL Science 294:2317-2323(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-21608551; PubMed-11743194; Miller N., Blanchard M.,
 RA Goodner B., Hinkle G., Gattung S., Cao Y., Askenzai M., Mullin L.,
 RA Houmelo B., Goldman B.S., Cao Y., Askenzai M., Halling C., Mullin L.,
 RA Houmelo B., Goldman B.S., Cao Y., Askenzai M., Halling C., Mullin L.,
 RA Wollam C., Allinger M., Doughty D., Scott C., Lappas C., Markelz B.,
 RA Plangan C., Crowell C., Gursen J., Lomo C., Sear C., Strud G.,
 RA Cielo C., Slater S.,
 RT "Genome sequence of the plant pathogen and biotechnology agent
 RT *Agrobacterium tumefaciens* C58."
 RL Science 294:2323-2328(2001).
 DR EMBL: AE009322; AAL44721.1;
 DR EMBL: AE008292; AAK89506.1;
 KW Complete proteome.
 SQ SEQUENCE 576 AA; 61722 MW; E0299479A952581F CRC64;
 Query Match 46.6%; Score 159.5; DB 16; Length 576;
 Best Local Similarity 50.7%; Pred. No. 3.3e-08;
 Matches 35; Conservative 13; Mismatches 20; Indels 1; Gaps 1;
 OY 1 EGEIPAPLAGVSKILVKEGDTYVAGQTVLVLEAKMKEINAPTDGKVEKLVKENDAV 60
 DB 507 EGEITAVSGTISGTFKVDGDETVSEGDLLVMEAMKMETQVAVRAGV-RLIVKEGDYL 565
 OY 61 QGGGGLIKI 69
 DB 566 QAGATLIDI 574
 RESULT 16
 O92H13 PRELIMINARY; PRT; 665 AA.
 AC O92H13;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DE 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Propionyl-CoA carboxylase alpha chain.
 GN PCCA OR RC0959.
 OS Rickettsia conorii.
 OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
 OC Rickettsiaceae; Rickettsiae; Rickettsia.
 OX NCBI_Taxid-781;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN-MALISH 7;
 RX MEDLINE-21442074; PubMed-11557893;
 RA Ogata H., Audic S., Renesto-Audiffren P., Fournier P.-E., Barde V.,
 RA Samson D., Roux V., Cossart P., Weissenbach J., Claverie J.-M.,
 RA Reout D.,
 RT "Mechanisms of evolution in Rickettsia conorii and R. prowazekii."
 RL Science 293:2093-2098(2001).
 DR EMBL: AE008649; AAL03497.1;
 DR InterPro: IPR001882; Biotin_attach.
 DR InterPro: IPR000089; Biotin_lipoyl.
 DR InterPro: IPR000901; CPase.
 DR Pfam: PF02785; Biotin_card_C_1.
 DR Pfam: PF00364; Biotin_lipoyl_1.
 DR Pfam: PF00289; CPase_L_Chain_1.
 DR Pfam: PF02786; CPase_L_D2; 1.
 DR PROSITE: PS00188; BIOTIN; UNKNOWN_1.
 DR PROSITE: PS00866; CPASE_1; UNKNOWN_1.

DR PROSITE: PS00867; CPASE_2; UNKNOWN_1.
 KW Complete proteome.
 SQ SEQUENCE 665 AA; 73320 MW; 3054C1E18E0C14A3 CRC64;
 Query Match 46.5%; Score 159; DB 16; Length 665;
 Best Local Similarity 50.0%; Pred. No. 4.3e-08;
 Matches 33; Conservative 12; Mismatches 21; Indels 0; Gaps 0;
 OY 3 EIPAPLAGVSKILVKEGDTYVAGQTVLVLEAKMKEINAPTDGKVEKLVKENDAV 62
 DB 598 ELGAPLSGQIAIKVKEGDTYVAGQTVLVLEAKMKEINAPTDGKVEKLVKENDAV 657
 OY 63 QGGGGLIKI 68
 DB 658 GQVILE 663
 RESULT 17
 O974R8 PRELIMINARY; PRT; 169 AA.
 AC O974R8;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DE 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Putative biotin carboxyl carrier protein.
 GN ST0592.
 OS Sulfolobus tokodaii.
 OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
 OC Sulfolobus.
 OX NCBI_Taxid-111955;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN-JCM 10545 / 7;
 RX PubMed-11572479;
 RA Kawarabayashi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,
 RA Sekine M., Baba S.-I., Anai A., Kosugi H., Hosoyama A., Fukui S.,
 RA Negai Y., Nishijima K., Otsuka R., Nakazawa H., Takamiya M., Kato Y.,
 RA Yoshizawa T., Tanaka T., Kudo Y., Yamazaki J., Kushi N., Oguchi A.,
 RA Aoki K.-I., Masuda S., Yanagii M., Nishimura M., Yamagishi A.,
 RA Oshima T., Kikuchi H.,
 RT "Complete genome sequence of an aerobic thermophilic
 RT *Crenarchaeon*, *Sulfolobus tokodaii* strain 7."
 RL DNA Res. 8:123-140(2001).
 DR EMBL: AP000983; BAB5589.1;
 DR InterPro: IPR000089; Biotin_lipoyl.
 DR Pfam: PF00364; Biotin_lipoyl_1.
 KW Hypothetical protein: Complete proteome.
 SQ SEQUENCE 169 AA; 18839 MW; 4F1D96761F5BE3D CRC64;
 Query Match 46.2%; Score 158; DB 17; Length 169;
 Best Local Similarity 49.3%; Pred. No. 1.2e-08;
 Matches 34; Conservative 12; Mismatches 23; Indels 0; Gaps 0;
 OY 1 EGEIPAPLAGVSKILVKEGDTYVAGQTVLVLEAKMKEINAPTDGKVEKLVKENDAV 60
 DB 100 EGEITAVSGTISGTFKVDGDETVSEGDLLVMEAMKMETQVAVRAGV-RLIVKEGDYL 565
 OY 61 QGGGGLIKI 69
 DB 160 KKGDLILIT 168
 RESULT 18
 O9HTD1 PRELIMINARY; PRT; 607 AA.
 AC O9HTD1;
 DT 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 DE 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Probable transcarboxylase subunit.
 GN PA5435.
 OS Pseudomonas aeruginosa.
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;

OC Pseudomonas.
 OX NCBI_TaxID=287;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 15692 / PA01;
 RX MEDLINE=20437337; PubMed=10984043;
 RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Wattner P.,
 RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
 RA Gardner R.L., Goltzy L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
 RA Barclay L.L., Coulter S.N., Folger K.R., Kas A., Lardig K., Linn R.M.,
 RA Smith K.A., Spencer D.H., Mong G.K.-S., Wu Z., Paulsen I.T.,
 RA Reiter J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.,
 RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
 RT opportunistic pathogen.";
 RL Nature 406:959-964(2000).
 DR EMBL: AE004956; AAG08820.1; -
 DR HSSP: P02905; 3BDO.
 DR InterPro: IPR000089; Biotin_lipoyl.
 DR InterPro: IPR000891; HMG1-like.
 DR InterPro: IPR003379; Pyc_OADA.
 DR InterPro: IPR000634; S/T_dehydrtse.
 DR Pfam: PF00364; biotin_lipoyl; 1.
 DR Pfam: PF00682; HMG1-like; 1.
 DR Pfam: PF02436; Pyc_OADA; 1.
 DR TIGRFAMS: TIGR01108; oada; 1.
 DR PROSITE: PS00165; DEHYDRATASE_SER_THR; UNKNOWN_1.
 DR Complete proteome.
 SQ SEQUENCE 607 AA; 66095 MW; 232AB0E9B935E010 CRC64;

Query Match 46.2%; Score 158; DB 16; Length 607;
 Best Local Similarity 48.5%; Pred. No. 4.9e-08;
 Matches 33; Conservative 12; Mismatches 23; Indels 0; Gaps 0;

OY 2 GEIPAPLACTVSKIVKESGDTVAGQTVLYEAKMKETEINAPTDGKVEKLYKERDVG 61
 DB 538 GHVSTMPKENVLYLVKEDSVKAGAVILTEAKMKETEVQAGIAGTVAIHAKKDRVN 597
 OY 62 GGQGLIKI 69
 DB 598 PGEILKEI 605

RESULT 19
 ID 092CWI PRELIMINARY; PRT; 1146 AA.
 AC 092CWI;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE PycA protein.
 GN PycA OR L1N1060.
 OS Listeria innocua.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
 OC Listeriaceae; Listeria.
 OX NCBI_TaxID=1642;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CLIP 11262 / SEROVAR 6A;
 RX PubMed=11679669;
 RA Glaser P., Frangeul L., Buchrieser C., Rusnlok C., Amend A.,
 RA Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,
 RA Charbit A., Chetouani F., Couve E., de Darvar A., Dehoux P.,
 RA Domain E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
 RA Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,
 RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
 RA Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kurapkai G.,
 RA Nordstedt G., Novella S., de Pablo B., Perez-Diaz J.-C., Purcell R.,
 RA Remmel B., Rose M., Schueter T., Simoes N., Tierrez A.,
 RA Vazquez-Boland J.-A., Voss H., Wehlund J., Cossart P.,
 RT "Comparative genomics of Listeria species.";
 RL Science 294:849-852(2001).
 DR EMBL: AL559167; CAC96291.1; -

DR L1stIlist; L1N01060; -
 DR InterPro: IPR000089; Biotin_lipoyl.
 DR InterPro: IPR000901; CPSase.
 DR InterPro: IPR000891; HMG1-like.
 DR InterPro: IPR003379; Pyc_OADA.
 DR Pfam: PF02785; Biotin_carb_C; 1.
 DR Pfam: PF00364; biotin_lipoyl; 1.
 DR Pfam: PF00289; CPSase_L_chain; 1.
 DR Pfam: PF02786; CPSase_L_D2; 1.
 DR Pfam: PF00682; HMG1-like; 1.
 DR Pfam: PF02436; Pyc_OADA; 1.
 DR TIGRFAMS: TIGR01235; pyrUV_carbox; 1.
 DR PROSITE: PS00866; CPSase_1; UNKNOWN_1.
 DR PROSITE: PS00867; CPSase_2; UNKNOWN_1.
 DR Complete proteome.
 SQ SEQUENCE 1146 AA; 128035 MW; 3E23FFB4A289C60F CRC64;

Query Match 45.9%; Score 157; DB 16; Length 1146;
 Best Local Similarity 43.9%; Pred. No. 1.3e-07;
 Matches 29; Conservative 17; Mismatches 20; Indels 0; Gaps 0;

OY 4 IPAPLAGTIVSKIVKESGDTVAGQTVLYEAKMKETEINAPTDGKVEKLYKERDVG 63
 DB 1078 VCATMGSVIYVVKKGDSVKKGPDLITAEAKMKETTIOAPPDGEVSSITVSDGDTISG 1137
 OY 64 GGLIKI 69
 DB 1138 DLIEV 1143

RESULT 20
 ID 08Y846 PRELIMINARY; PRT; 1146 AA.
 AC 08Y846;
 DT 01-MAR-2002 (TREMBLrel. 20, Created)
 DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE PycA protein.
 GN PycA OR LMO1072.
 GN PycA OR LMO1072.
 OS Listeria monocytogenes.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
 OC Listeriaceae; Listeria.
 OX NCBI_TaxID=1639;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-EGD-E / SEROVAR 1/2A;
 RX MEDLINE=21537279; PubMed=11679669;
 RA Glaser P., Frangeul L., Buchrieser C., Rusnlok C., Amend A.,
 RA Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,
 RA Charbit A., Chetouani F., Couve E., de Darvar A., Dehoux P.,
 RA Domain E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
 RA Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,
 RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
 RA Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kurapkai G.,
 RA Madueno E., Maitouram A., Mata Vicente J., Ng E., Nedjati H.,
 RA Nordstedt G., Novella S., de Pablo B., Perez-Diaz J.-C., Purcell R.,
 RA Remmel B., Rose M., Schueter T., Simoes N., Tierrez A.,
 RA Vazquez-Boland J.-A., Voss H., Wehlund J., Cossart P.,
 RT "Comparative genomics of Listeria species.";
 RL Science 294:849-852(2001).
 DR EMBL: AL559177; CAC99150.1; -

DR L1stIlist; LMO1072; -
 DR InterPro: IPR000089; Biotin_lipoyl.
 DR InterPro: IPR000901; CPSase.
 DR InterPro: IPR000891; HMG1-like.
 DR InterPro: IPR003379; Pyc_OADA.
 DR Pfam: PF02785; Biotin_carb_C; 1.
 DR Pfam: PF00364; biotin_lipoyl; 1.
 DR Pfam: PF00289; CPSase_L_chain; 1.
 DR Pfam: PF02786; CPSase_L_D2; 1.
 DR Pfam: PF00682; HMG1-like; 1.
 DR Pfam: PF02436; Pyc_OADA; 1.
 DR TIGRFAMS: TIGR01235; pyrUV_carbox; 1.

DR * PROSITE: PS00866; CPSASE_1; 1.
DR PROSITE: PS00867; CPSASE_2; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 1146 AA; 128049 MW; A062F88C9A092B6F CRC64;
Query Match 45.9%; Score 157; DB 16; Length 1146;
Best Local Similarity 43.9%; Pred. No. 1.3e-07;
Matches 29; Conservative 17; Mismatches 20; Indels 0; Gaps 0;
QY 4 IPAPLAGTAVSKILVKGDTVAKGQTVLVLEAMKMETEINAPDGVKVERVYKERVADAVOGG 63
DB 1078 VQATMTGVSIVQVYKKGDSVYKKGDELITLTKMKMETITQAPFDGVSIVYSDGTTISG 1137
QY 64 QGLIKI 69
DB 1138 DLTLEV 1143
RESULT 21
QY 09K9M0 PRELIMINARY; PRT; 1150 AA.
AC 09K9M0;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Pyruvate carboxylase.
GN PYCA OR BH2625.
OS Bacillus halodurans.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacilliales;
OC Bacillaceae; Bacillus.
CX NCBI_TaxID=86665;
KM NCBI_TaxID=86665;
SQ SEQUENCE FROM N.A.
RA STRAIN-C-125 / JCM 9153;
RX MEDLINE=20512582; PubMed=11058132;
RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA Fujii F., Hirama C., Nakamura Y., Ogasawara N., Kunara S.,
RA Horikoshi K.;
RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
RT halodurans and genomic sequence comparison with Bacillus subtilis.";
RL Nucleic Acids Res. 28:4317-4331(2000).
DR EMBL: AP001516; BAB06344.1; -
DR HSSP: P24182; 1BKC.
DR InterPro: IPR000089; Biotin_lipoyl.
DR InterPro: IPR000901; CPSase.
DR InterPro: IPR000891; HMGL-like.
DR InterPro: IPR003379; PYC_OADA.
DR Pfam: PF02785; Biotin_carb_C; 1.
DR Pfam: PF00364; Biotin_lipoyl; 1.
DR Pfam: PF00289; CPSase_L-chain; 1.
DR Pfam: PF02786; CPSase_L-D2; 1.
DR Pfam: PF00682; HMGL-like; 1.
DR Pfam: PF02436; PYC_OADA; 1.
DR TIGRFAMs: TIGR01235; pyruv_carbox; 1.
DR PROSITE: PS00867; CPSASE_2; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 1150 AA; 128858 MW; ED6788BE8A9F3BD4 CRC64;
Query Match 45.9%; Score 157; DB 16; Length 1150;
Best Local Similarity 50.0%; Pred. No. 1.3e-07;
Matches 33; Conservative 13; Mismatches 20; Indels 0; Gaps 0;
QY 4 IPAPLAGTAVSKILVKGDTVAKGQTVLVLEAMKMETEINAPDGVKVERVYKERVADAVOGG 63
DB 1082 IGASMPGTVVVALVEKGVKVGSHLMTLTKMKMETITQAPFDGVSIVYSDGTTISG 1141
QY 64 QGLIKI 69
DB 1142 DLTLEV 1147
RESULT 22
QY 099ZL6

ID 099ZL6 PRELIMINARY; PRT; 116 AA.
AC 099ZL6;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Putative methylmalonyl-CoA decarboxylase, gamma-subunit
DE (EC 4.1.1.41).
GN SPY1176.
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
OC Streptococcaceae; Streptococcus.
CX NCBI_TaxID=1314;
KM NCBI_TaxID=1314;
SQ SEQUENCE FROM N.A.
RA STRAIN-SF370 / ATCC 700294 / SEROTYPE M1;
RX MEDLINE=21192684; PubMed=11296296;
RA Ferretti J.J., Meshan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K.,
RA Primeaux C., Sezale S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P.,
RA Qian Y., Jia H.G., Najjar F.Z., Ren Q., Zhu H., Song L., White J.,
RA Yuan X., Clifton S.W., Roe B.A., McLaughlin R.;
RT "Complete genome sequence of an M1 strain of Streptococcus pyogenes.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).
CC -1-COFACITOR: BIOTIN (BY SIMILARITY).
DR EMBL: AE006558; AAK34043.1; -
DR HSSP: P02905; 1BDO.
DR InterPro: IPR001882; Biotin_attach.
DR InterPro: IPR000089; Biotin_lipoyl.
DR Pfam: PF00364; Biotin_lipoyl; 1.
DR PROSITE: PS00188; BIOTIN; 1.
KM Biotin_Lyase; Complete proteome.
SQ SEQUENCE 116 AA; 12288 MW; A245219AC595BFEC CRC64;
Query Match 45.6%; Score 156; DB 16; Length 116;
Best Local Similarity 52.3%; Pred. No. 1.2e-08;
Matches 34; Conservative 7; Mismatches 24; Indels 0; Gaps 0;
QY 3 EIPAPLAGTAVSKILVKGDTVAKGQTVLVLEAMKMETEINAPDGVKVERVYKERVADAVOGG 62
DB 50 QVKAWSGTSLIFATEKGAVKKGSAVLVLEAMKMETEINAPDGLVSKIHVAVQYTES 109
QY 63 QGLIKI 67
DB 110 EQVLI 114
RESULT 23
QY 08ROL2 PRELIMINARY; PRT; 1139 AA.
AC 08ROL2;
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Pyruvate carboxylase.
GN PYC.
OS Corynebacterium efficiens.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacteriaceae; Corynebacteriaceae;
OC Corynebacterium.
CX NCBI_TaxID=152794;
KM NCBI_TaxID=152794;
SQ SEQUENCE FROM N.A.
RA Akiyoshi N., Nonaka G., Kimura E., Kawahara Y., Sugimoto S.;
RT "Corynebacterium efficiens pyruvate carboxylase (pyc) gene, complete
RT CDS.";
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB083299; BAB88903.1; -
SQ SEQUENCE 1139 AA; 123068 MW; BA7023134519FAAA CRC64;
Query Match 45.5%; Score 155.5; DB 2; Length 1139;
Best Local Similarity 43.5%; Pred. No. 1.8e-07;
Matches 30; Conservative 15; Mismatches 23; Indels 1; Gaps 1;
QY 1 EGEIPAPLAGTAVSKILVKGDTVAKGQTVLVLEAMKMETEINAPDGVKVERVYKERVADAVOGG 60

```

Db      1071 KGHVAAFPAGVVT-VTVAESDEIKAGDAVAIIEMKMEATITAPDVGDIVVVPAATKY 1129
QY      61 QGGGGLIKI 69
Db      1130 EGGDLIVV 1138

RESULT 24
054587
AC      054587; PRELIMINARY; PRT; 1140 AA.
ID      054587;
DT      01-JUN-1998 (TREMblrel. 06, Created)
DT      01-JUN-1998 (TREMblrel. 06, Last sequence update)
DT      01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE      Pyruvate carboxylase.
GN      Pyc.
OS      Corynebacterium glutamicum (Brevibacterium flavum).
OC      Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC      Actinomycetales; Corynebacterineae; Corynebacteriaceae;
OC      Corynebacterium.
OX      NCBI_TaxID=1718;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=21253;
RA      Kofitis M.A.G., Ramamoorthi R., Pine W.A., Sinsky A.J.,
RA      Stephanopoulos G.,
RL      Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
RN      [2]
RP      SEQUENCE FROM N.A.
RC      STRAIN=ATCC 13032;
RA      Peters-Wendisch P.G., Kreutzer C., Kallowski J., Patek M., Sahn H.,
RA      Elkmann B.J.;
RL      Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
CC      -1- COFACTOR: BIOTIN (BY SIMILARITY).
DR      EMBL; AF038548; AAB92588.1; -.
DR      EMBL; Y09548; CA70739.1; -.
DR      HSSP; P24182; IBC.
DR      InterPro: IPR001882; Biotin_attach.
DR      InterPro: IPR000089; Biotin_lipoyl.
DR      InterPro: IPR000901; CPSase.
DR      InterPro: IPR000891; HMGL-like.
DR      InterPro: IPR003379; Pyc_OADA.
DR      Pfam; PF02785; Biotin_card_C; 1.
DR      Pfam; PF00364; Biotin_lipoyl; 1.
DR      Pfam; PF00289; CPSase_L-chain; 1.
DR      Pfam; PF02786; CPSase_L_D2; 1.
DR      Pfam; PF00682; HMGL-like; 1.
DR      Pfam; PF02436; Pyc_OADA; 1.
DR      TIGRfams; TIGR01235; pyruv_carbox; 1.
DR      PROSITE; PS00188; BIOTIN; 1.
DR      PROSITE; PS00867; CPSASE_2; UNKNOWN_1.
KW      Biotin; ligase; Pyruvate.
SQ      SEQUENCE 1140 AA; 123102 MW; 2A6D4B4ED2FEB531 CRC64;

Query Match      45.5%; Score 155.5; DB 2; Length 1140;
Best Local Similarity 44.9%; Pred. No. 1.8e-07;
Matches 31; Conservative 14; Mismatches 23; Indels 1; Gaps 1;

Db      1072 KGHVAAFPAGVVT-VTVAESDEIKAGDAVAIIEMKMEATITASVDSKIDRVVPAATKY 1130
QY      61 QGGGGLIKI 69
Db      1131 EGGDLIVV 1139

RESULT 25
08RAJ2
ID      08RAJ2; PRELIMINARY; PRT; 135 AA.
AC      08RAJ2;
DT      01-JUN-2002 (TREMblrel. 21, Created)

```

```

DT      01-JUN-2002 (TREMblrel. 21, Last sequence update)
DE      01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE      Biotin carboxyl carrier protein.
GN      ACCB OR TFE122.
OS      Thermoaerobacter tengcongensis.
OC      Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC      Thermoaerobacteriales; Thermoaerobacteriaceae; Thermoaerobacter.
OX      NCBI_TaxID=119072;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=MB47 / JCM11007;
RX      MEDLINE=21992816; PubMed=11997336;
RA      Bao Q., Tian Y., Li W., Xu Z., Xuan Z., Hu S., Dong W., Yang J.,
RA      Chen Y., Xue Y., Xu Y., Lai X., Huang L., Dong X., Ma Y., Ling L.,
RA      Tan H., Chen R., Wang J., Yu J., Yang H.;
RT      "A complete sequence of T. tengcongensis genome.";
RL      Genome Res. 12:689-700(2002).
DR      EMBL; AE013084; AAM24452.1; -.
KW      Complete proteome.
SQ      SEQUENCE 135 AA; 15195 MW; 9C2BF717C008800A CRC64;

Query Match      44.7%; Score 153; DB 16; Length 135;
Best Local Similarity 47.8%; Pred. No. 2.9e-08;
Matches 33; Conservative 11; Mismatches 23; Indels 2; Gaps 1;

QY      59 AVGGGGLI 67
Db      64 EGEKTIAPMPGRTIVYRVKESGDKVRKGDVITLLEMKMENEIMAPENGITVSVNAKDD 123
Db      124 TVNRGDIV 132

RESULT 26
067544
ID      067544; PRELIMINARY; PRT; 620 AA.
AC      067544;
DT      01-AUG-1998 (TREMblrel. 07, Created)
DT      01-AUG-1998 (TREMblrel. 07, Last sequence update)
DT      01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE      Oxalacetate decarboxylase alpha chain.
GN      OADA OR AO_1614.
OS      Aquifex aeolicus.
OC      Bacteria; Aquificae; Aquificales; Aquificaceae; Aquifex.
OX      NCBI_TaxID=63363;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=VF5;
RX      MEDLINE=98196666; PubMed=9537320;
RA      Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
RA      Graham D.E., Overbeek R., Sneed M.A., Keller M., Anjey M., Huber R.,
RA      Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
RT      "The complete genome of the hyperthermophilic bacterium Aquifex
RT      aeolicus.";
RL      Nature 392:353-358(1998).
CC      -1- FUNCTION: TUBULIN IS THE MAJOR CONSTITUENT OF MICROTUBULES. IT
CC      BINDS TWO MOLES OF GTP, ONE AT AN EXCHANGEABLE SITE ON THE BETA
CC      CHAIN AND ONE AT A NONEXCHANGEABLE SITE ON THE ALPHA-CHAIN (BY
CC      SIMILARITY).
CC      -1- COFACTOR: BIOTIN (BY SIMILARITY).
CC      -1- SIMILARITY: BELONGS TO THE TUBULIN FAMILY.
DR      EMBL; AE000747; AAC07497.1; -.
DR      HSSP; P02905; IBD0.
DR      InterPro: IPR002453; Beta_tubulin.
DR      InterPro: IPR001882; Biotin_attach.
DR      InterPro: IPR000089; Biotin_lipoyl.
DR      InterPro: IPR000891; HMGL-like.
DR      InterPro: IPR003379; Pyc_OADA.
DR      Pfam; PF00364; Biotin_lipoyl; 1.
DR      Pfam; PF00682; HMGL-like; 1.
DR      Pfam; PF02436; Pyc_OADA; 1.
DR      TIGRfams; TIGR01108; oada; 1.

```


DR PROSITE: PS00188; BIOTIN; 1.
DR PROSITE: PS00228; TUBULIN_B; AUTOREG; 1.
KW Biotin; GTP-binding; Microtubules; Complete proteome.
SQ SEQUENCE 620 AA; 70416 MW; 4F306D48794AE859 CRC64;

Query Match 44.7%; Score 153; DB 16; Length 620;
Best Local Similarity 45.6%; Pred. No. 1.6e-07;
Matches 31; Conservative 14; Mismatches 23; Indels 0; Gaps 0;

QY 2 GEIPAPLAGVSKILVEGDTVAKGQTVLVLEAMKMETEINAPTDGKVEKYLKERDAVQ 61
DB 545 GDVISPITGKVNKNKAVGVDEKGDVLLVLEAMKMETEINAPTDGKVEKYLKERDAVQ 604
QY 62 GGCGGLIKI 69
DB 605 PDEVLIRI 612

RESULT 27
052603 PRELIMINARY; PRT; 167 AA.
AC 052603;
DT 01-JUN-1998 (TREMBlrel. 06, Created)
DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Biotin carboxyl carrier protein.
GN ACCB.
OS Sulfolobus metallicus.
OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
OC Sulfolobus.
RN NCBI_TaxID=47303;
RX MEDLINE-20059326; PubMed=10591844;
RC STRAIN-LM;
RP SEQUENCE FROM N.A.
RA Burton N.P., Williams T.D., Norris P.R.;
RT "Carboxylase genes of Sulfolobus metallicus";
RL Arch. Microbiol. 172:349-353(1999).
DR HSSP; P02905; IBD0.
DR InterPro: IPR000089; Biotin_lipoyl.
DR Pfam: PF00364; biotin_lipoyl.
SQ SEQUENCE 167 AA; 18580 MW; F05A04F38BC55B09 CRC64;

Query Match 44.4%; Score 152; DB 1; Length 167;
Best Local Similarity 50.7%; Pred. No. 4.7e-08;
Matches 34; Conservative 9; Mismatches 24; Indels 0; Gaps 0;

QY 1 EGEIPAPLAGVSKILVEGDTVAKGQTVLVLEAMKMETEINAPTDGKVEKYLKERDAV 60
DB 98 EGEIPAPLAGVSKILVEGDTVAKGQTVLVLEAMKMETEINAPTDGKVEKYLKERDAV 157
QY 61 GGCGGLIKI 67
DB 158 KKGDLLV 164

RESULT 28
040121 PRELIMINARY; PRT; 436 AA.
AC 040121;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Biotin-containing subunit of methylcrotonyl-CoA carboxylase
DE (fragment).
OS Lycopersicon esculentum (Tomato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
RN NCBI_TaxID=4081;
RP SEQUENCE FROM N.A.

RX MEDLINE-94216274; PubMed=8163472;
RA Wang X., Wurttele E.S., Keller G., McKean A.L., Nikolau B.J.;
RT "Molecular cloning of cDNAs and genes coding for beta-methylcrotonyl-
RT CoA carboxylase of tomato."
RL J. Biol. Chem. 269:11760-11768(1994).
CC -|- COFACTOR: BIOTIN (BY SIMILARITY).
DR EMBL; U07745; AAA19157.1; -.
DR HSSP; P24182; IDV1.
DR InterPro: IPR001882; Biotin_attach.
DR InterPro: IPR000089; Biotin_lipoyl.
DR InterPro: IPR000901; CPase.
DR Pfam; PF02785; Biotin_carb_C; 1.
DR Pfam; PF00364; Biotin_lipoyl; 1.
DR Pfam; PF02786; CPase_L_D2; 1.
DR PROSITE; PS00188; BIOTIN; 1.
DR PROSITE; PS00867; CPASE_2; UNKNOWN_1.
KW Biotin.
FT NON_TER
SQ SEQUENCE 1 1
48414 MW; A9E4110E781E4944 CRC64;

Query Match 44.4%; Score 152; DB 10; Length 436;
Best Local Similarity 51.6%; Pred. No. 1.4e-07;
Matches 32; Conservative 11; Mismatches 19; Indels 0; Gaps 0;

QY 2 GEIPAPLAGVSKILVEGDTVAKGQTVLVLEAMKMETEINAPTDGKVEKYLKERDAVQ 61
DB 367 GTVIAFPAAGLVKVLVKGDEKGVLPVLEAMKMETEINAPTDGKVEKYLKERDAVQ 426
QY 62 GG 63
DB 427 DG 428

RESULT 29
099YD8 PRELIMINARY; PRT; 166 AA.
AC 099YD8;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Putative biotin carboxyl carrier protein.
GN ACBB OR SPY1747.
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
OC Streptococcaceae; Streptococcus.
RN NCBI_TaxID=1314;
RP SEQUENCE FROM N.A.
RC STRAIN-SF370 / SEROTYPE M1;
RX MEDLINE-21192684; PubMed=11296296;
RA Ferretti J.J., McShan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K.,
RA Primeaux C., Sezate S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P.,
RA Qian Y., Jia H.G., Najjar F.Z., Ren O., Zhu H., Song L., White J.,
RA Yuan X., Clifton S.W., Roe B.A., McLaughlin R.;
RT "Complete genome sequence of an M1 strain of Streptococcus pyogenes";
RT Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).
DR EMBL; AE006603; AAK34491.1; -.
DR HSSP; P02905; IBD0.
DR InterPro: IPR001249; ACCOA_biotinCC.
DR InterPro: IPR001882; Biotin_attach.
DR InterPro: IPR000089; Biotin_lipoyl.
DR Pfam; PF00364; Biotin_lipoyl; 1.
DR PRINTS; PRO1071; ACOAB10TINC.
DR TIGRFAMs; TIGR00531; BCCP; 1.
DR PROSITE; PS00188; BIOTIN; 1.
KW Complete proteome.
SQ SEQUENCE 166 AA; 17901 MW; 46EBCB229A2C154D CRC64;

Query Match 44.2%; Score 151; DB 16; Length 166;
Best Local Similarity 41.6%; Pred. No. 6e-08;
Matches 32; Conservative 15; Mismatches 22; Indels 8; Gaps 2;

QY 1 EGEIPAPLAGV-----SKILVEGDTVAKGQTVLVLEAMKMETEINAPTDGKVEK 52

```

Db      89 EGDVSPPLGVAVYLAASDPKPPVAVGDTPVKGQTLVITEAKVMNEVPAPCDGVITEI 148
QY      53 LVKERDVOGGGGLIKI 69
Db      149 LVSNEDVIERGGGLVRI 165

```

RESULT 30

O9KDS8

PRELIMINARY; PRT; 70 AA.

ID O9KDS8

AC O9KDS8: 01-OCT-2000 (TReMBLrel. 15, Created)

DT 01-OCT-2000 (TReMBLrel. 15, last sequence update)

DT 01-MAR-2002 (TReMBLrel. 20, last annotation update)

DE Hypothetical protein BH1133.

GN BH1133.

OS Bacillus halodurans.

OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacilliales;

OC Bacillaceae; Bacillus.

OX NCBI_TaxID=86665;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C-125 / JCM 9153;

RX MEDLINE=20512582; PubMed=11058132;

RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,

RA Fuji F., Hirata C., Nakamura Y., Ogasawara N., Kuhara S.,

RA Horikoshi K.;

RT *Complete genome sequence of the alkaliphilic bacterium Bacillus

RT halodurans and genomic sequence comparison with Bacillus subtilis."

RT Nucleic Acids Res. 28:4317-4331(2000).

RL EMBL: AF001511; BAB04852.1; -.

DR HSSP: P10802; IITD.

DR InterPro: IPR000089; Biotin_lipoY1.

DR Pfam: PF00364; biotin_lipoY1; 1.

KM Hypothetical protein; Complete proteome.

SQ SEQUENCE 70 AA: 7563 MW: 6A6AC14B14B5F7AA CRC64;

Query Match

Best Local Similarity 43.9%; Score 150; DB 16; Length 70;

Matches 31; Conservative 15; Mismatches 21; Indels 0; Gaps 0;

```

QY      3 ETPAPLAGVSKLVKEGDTVAKGOTVLYLEAKMETETINAPIDGKVEKVLKERDVOG 62
      1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      2 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      3 KIETNAGNWKILVKEGQVEAGVEAILESKMEIPVEAASGTVKSVLKQEGEIDE 62
QY      63 GGLIKI 69
      1 : : : : :
Db      63 GGLIEL 69

```

Search completed: May 1, 2003, 07:52:25
 Job time : 39.2812 secs